

UNIVERSIDAD TÉCNICA DE AMBATO



FACULTAD DE SISTEMAS, ELECTRONICA E INDUSTRIAL

MAESTRIA EN MATEMATICA APLICADA

Tema: “MODELO ESTADÍSTICO PARA VALORAR LAS LÍNEAS DE TRANSPORTE PÚBLICO DE PASAJEROS EN LA CIUDAD DE RIOBAMBA”

Trabajo de titulación previo a la obtención del grado académico de
Magister en Matemática Aplicada
Modalidad de titulación proyecto de desarrollo

Autor: Fabián Patricio Londo Yachambáy

Director: Dr. Manuel Antonio Meneses Freire. PhD.

Ambato – Ecuador

Año 2021

APROBACIÓN DEL TRABAJO DE TITULACIÓN

A la Unidad Académica de Titulación de la Facultad de Ingeniería en Sistemas, Electrónica e Industrial. El Tribunal receptor de la Defensa del Trabajo de Titulación presidido por la Ing. Elsa Pilar Urrutia Urrutia Mg., e integrado por los señores: Lic. Daniel Morocho Lara, Mg. y el Lic. José Nicolas Torrealba Mg, designados por la Unidad Académica de Titulación de Posgrados de la Facultad de Ingeniería en Sistemas Electrónica e Industrial de la Universidad Técnica de Ambato, para receptor el Trabajo de Titulación con el tema: “Modelo estadístico para valorar las líneas de transporte público de pasajeros en la ciudad de Riobamba”, elaborado y presentado por el Ing. Fabián Patricio Londo Yachambáy, para optar por el Grado Académico de Magister en Matemática Aplicada; una vez escuchada la defensa oral del Trabajo de Titulación el Tribunal aprueba y remite el trabajo para uso y custodia en las bibliotecas de la Universidad Técnica de Ambato.



Ing. Elsa Pilar Urrutia Urrutia, Mg.
Presidente y Miembro del Tribunal de Defensa



Lic. Daniel Morocho Lara, Mg.
Miembro del Tribunal de Defensa



Lic. José Nicolas Torrealba, Mg.
Miembro del Tribunal de Defensa

AUTORÍA DEL TRABAJO DE TITULACIÓN

La responsabilidad de las opiniones, comentarios y críticas emitidas en el trabajo de titulación presentado con el tema: “Modelo estadístico para valorar las líneas de transporte público de pasajeros en la ciudad de Riobamba”, le corresponde exclusivamente a: Ing. Fabián Patricio Londo Yachambáy, autor bajo la dirección de Dr. Manuel Antonio Meneses Freire, PhD, director del trabajo de investigación; y el patrimonio intelectual pertenece a la Universidad Técnica de Ambato.



Ing. Fabian Patricio Londo Yachambáy
AUTOR



Dr. Manuel Antonio Meneses Freire, PhD.
DIRECTOR

DERECHOS DE AUTOR

Autorizo a la Universidad Técnica de Ambato, para que el Trabajo de Titulación, sirva como un documento disponible para su lectura, consulta y procesos de investigación, según las normas de la Institución.

Cedo los Derechos de mi Trabajo de Titulación, con fines de difusión pública, además apruebo la reproducción de este, dentro de las regulaciones de la Universidad Técnica de Ambato.



Ing. Fabian Patricio Londo Yachambáy
c.c. 0603016114

ÍNDICE GENERAL

CONTENIDO	Pag.
PORTADA.....	i
APROBACIÓN DEL TRABAJO DE TITULACIÓN	ii
AUTORÍA DEL TRABAJO DE TITULACIÓN	iii
DERECHOS DE AUTOR	iv
ÍNDICE DE TABLAS	vii
ÍNDICE DE FIGURAS.....	viii
AGRADECIMIENTO	ix
DEDICATORIA	x
RESUMEN EJECUTIVO.....	xi
EXECUTIVE ABSTRACT	xii
CAPÍTULO I	1
PROBLEMA DE INVESTIGACIÓN	1
1.1 Introducción.....	1
1.2 Justificación	2
1.3 Objetivos.....	3
1.3.1 General.....	3
1.3.2 Específicos	3
CAPÍTULO II.....	4
ANTECEDENTES INVESTIGATIVOS	4
CAPÍTULO III.....	12
MARCO METODOLÓGICO.....	12
3.1 Ubicación	12
3.2 Equipos y materiales	12
3.3 Nivel o tipo de investigación	13
3.3.1 Investigación Exploratoria.....	13
3.3.2 Investigación Descriptiva.....	13
3.3.3 Investigación Explicativa.....	13
3.3.4 Investigación Correlacional	13
3.4 Prueba de Hipótesis.....	13
3.5 Población o muestra.....	14
3.6 Recolección de Información	15
3.7 Procesamiento de la información y análisis estadístico	15
3.8 Variables respuesta o resultados alcanzados.....	16
CAPÍTULO IV.....	17
RESULTADOS Y DISCUSIÓN	17
4.1 Análisis e interpretación.....	17

4.1.1	Imputación de datos faltantes.....	17
4.1.2	Análisis descriptivo.....	18
4.1.3	Análisis descriptivo por pares de variables.....	26
4.1.4	Contraste de independencia Chi-cuadrado.....	30
4.1.5	Análisis de modelo de predicción	32
4.1.6	Validación del modelo	36
4.1.7	Presentación del modelo final.....	43
4.1.8	Contraste de razón de verosimilitudes del modelo multinomial	46
4.1.9	Análisis de la matriz pasajeros e Infraestructura de las vías (matriz 2)	47
4.1.10	Análisis descriptivo de Variables cualitativas.....	49
4.1.11	Análisis descriptivo de variables cuantitativas	55
4.1.12	Análisis del supuesto de Normalidad multivariada.....	58
4.1.13	Análisis y componentes principales y análisis factorial.....	60
4.1.14	Análisis de correspondencia.....	66
CAPÍTULO V.....		69
5.1	Conclusiones.....	69
5.2	Recomendaciones	70
5.3	BIBLIOGRAFÍA	71
5.4	ANEXOS	73

ÍNDICE DE TABLAS

Tabla 4-1: Valores faltantes por variable.....	17
Tabla 4-2: Valores imputados por variable.....	17
Tabla 4-3: Frecuencias de la variable Valoración de transporte	18
Tabla 4-4: Frecuencias de la variable Inconvenientes de Aglomeración.....	20
Tabla 4-5: Frecuencias de la variable Tiempo de origen a destino.....	21
Tabla 4-6: Frecuencias de la variable Razón de uso del transporte público	22
Tabla 4-7: Frecuencias de la variable Trato por conductor.....	23
Tabla 4-8: Frecuencias de la variable Kit de seguridad	25
Tabla 4-9: Test Chi-cuadrado de independencia.....	32
Tabla 4-10: Variables dummy según categorías de cada variable	33
Tabla 4-11: Estadísticos de Wald.....	34
Tabla 4-12: Análisis de aporte al modelo de cada variable dummy	35
Tabla 4-13: Tasa de clasificación correcta e incorrecta del modelo 1	37
Tabla 4-14: Tasa de clasificación correcta e incorrecta del modelo 2	38
Tabla 4-15: Tabla de resumen de los modelos elegidos	40
Tabla 4-16: Cálculo de las ODD ratios del modelo final.....	42
Tabla 4-17: Coeficientes del modelo final.....	43
Tabla 4-18: Anova del contraste de razón de verosimilitudes	46
Tabla 4-19: Proporción de valores faltantes por variable	47
Tabla 4-20: Imputación de datos faltantes	48
Tabla 4-21: Frecuencias de la variable Tipo de superficie	49
Tabla 4-22: Frecuencias de la variable Tipo de señalización vial.....	51
Tabla 4-23: Frecuencias de la variable Estado de superficie de rodadura	52
Tabla 4-24: Frecuencias de la variable Existencia de parada	53
Tabla 4-25: Frecuencias de la variable Señalización de parada.....	54
Tabla 4-26: Media aritmética de variables cuantitativas.....	56
Tabla 4-27: Cuartiles por variable	56
Tabla 4-28: Test de normalidad univariada	58
Tabla 4-29: Test de normalidad multivariada de Royston.....	60
Tabla 4-30: Matriz de correlaciones de las variables cuantitativas.....	61
Tabla 4-31: Coeficientes de cada componente	64
Tabla 4-32: Variabilidad explicada por cada componente.....	65

ÍNDICE DE FIGURAS

Figura 4-1: Frecuencia absoluta de la variable Valoración de transporte	19
Figura 4-2: Frecuencia absoluta de la variable Inconvenientes de Aglomeración.....	20
Figura 4-3: Frecuencias absoluta de la variable Tiempo de origen a destino	22
Figura 4-4: Frecuencias absoluta de la variable Razón de utilizar transporte.....	23
Figura 4-5: Frecuencias absoluta de la variable Trato por conductor	24
Figura 4-6: Frecuencias absoluta de la variable Kit de seguridad	25
Figura 4-7: Frecuencias cruzadas de Valoración de transporte e inconvenientes de aglomeración.....	26
Figura 4-8: Frecuencias cruzadas de Valoración de transporte y tiempo de origen a destino	27
Figura 4-9: Frecuencias cruzadas de Valoración de transporte y Razón de uso del transporte	28
Figura 4-10: Frecuencias cruzadas de Valoración de transporte y trato por conductor	29
Figura 4-11: Frecuencias cruzadas de Valoración de transporte y Kit de seguridad	30
Figura 4-12: Frecuencias absolutas de la variable Tipo de superficie	50
Figura 4-13: Frecuencias absolutas de la variable Tipo de señalización vial	51
Figura 4-14: Frecuencias absolutas de la variable Estado de superficie de rodadura	53
Figura 4-15: Frecuencias absolutas de la variable Existencia de parada	54
Figura 4-16: Frecuencias absolutas de la variable Señalización en parada.....	55
Figura 4-17: Histograma de variables cuantitativas.....	58
Figura 4-18: Q-Q plot de normalidad de las variables cuantitativas.....	60
Figura 4-19: Figura de correlación de las variables numéricas.....	62
Figura 4-20: Scree plot de factores y componentes a utilizar	64
Figura 4-21: Componente 1 vs. Componente 2	66
Figura 4-22: Porcentaje de variabilidad explicada por cada dimensión.....	67
Figura 4-23: Correspondencia multivariada de las categorías	68

AGRADECIMIENTO

Dios todo lo puede, por tal razón mi agradecimiento total a él, que me dio la sabiduría, la fuerza y la tenacidad para culminar con éxito este trabajo de investigación.

Gracias a mis padres porque debido a sus enseñanzas que siempre cultivaron en mí como la perseverancia y la constancia se pudo cumplir con el objetivo final.

También debo agradecer a mi esposa quien fue el pilar fundamental para colaborarme, apoyarme y apuntalarme en este sueño hecho realidad como es la culminación del trabajo final de grado.

DEDICATORIA

Este trabajo de investigación se la dedico a toda mi familia, pero de manera especial y con mucho amor y cariño a mi esposa quien me apoya en cada momento de mi vida, y sin duda también va dedicado con un cariño especial a mis dos adorados hijos Steven y Rafaela.

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE SISTEMAS, ELECTRÓNICA E INDUSTRIAL
CENTRO DE POSGRADOS
MAESTRIA EN MATEMÁTICA APLICADA

TEMA:

MODELO ESTADÍSTICO PARA VALORAR LAS LÍNEAS DE TRANSPORTE PÚBLICO DE PASAJEROS EN LA CIUDAD DE RIOBAMBA.

AUTOR: Ingeniero Fabian Patricio Londo Yachambáy

DIRECTOR: Dr. Manuel Antonio Meneses Freire PhD.

LÍNEA DE INVESTIGACIÓN

- Diseño, materiales y producción

FECHA: 27 de enero de 2021

RESUMEN EJECUTIVO

El transporte público de pasajeros es considerado como un servicio poco eficiente en el Ecuador y por ende en la ciudad de Riobamba, aunque se ha intentado por varias ocasiones alcanzar niveles de seguridad, comodidad y calidad altos es casi nada lo que se a podido mejorar. Ante esta problemática se trata aplicar un modelo estadístico que me permita valorar el transporte público de pasajeros, para lo cual se realizó un análisis en tres ejes que involucran a este medio de transporte como es Infraestructura, ascenso y descenso de pasajeros y calidad del servicio aplicando la técnica de encuesta y utilización un cuestionario para la obtención de información, con lo cual se obtuvo variables que ayudaron a construir mi modelo Logit Multinomial.

PALABRAS CLAVE: Modelo Estadístico, Movilidad, Politomica, Dummy, Logit Multinomial, Análisis de Componentes Principales, R-Macfaden, Kaiser Meyer, Cox Shell, Coeficiente de Wald.

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE SISTEMAS, ELECTRÓNICA E INDUSTRIAL
CENTRO DE POSGRADOS
MAESTRIA EN MATEMÁTICA APLICADA

TEMA:

MODELO ESTADÍSTICO PARA VALORAR LAS LÍNEAS DE TRANSPORTE PÚBLICO DE PASAJEROS EN LA CIUDAD DE RIOBAMBA.

AUTOR: Ingeniero Fabian Patricio Londo Yachambáy

DIRECTOR: Dr. Manuel Antonio Meneses Freire PhD.

LÍNEA DE INVESTIGACIÓN

- Diseño, materiales y producción

FECHA: 27 de enero de 2021

EXECUTIVE ABSTRACT

Public passenger transport is considered an inefficient service in Ecuador and therefore in the city of Riobamba, although there have been attempts on several occasions to achieve high levels of safety, comfort and quality, there is almost nothing that could be improved. Faced with this problem, it is about applying a statistical model that allows me to assess public passenger transport, for which an analysis was carried out in three axes that involve this means of transport, such as Infrastructure, ascent and descent of passengers and quality of service applying the survey technique and use of a questionnaire to obtain information, with which variables were obtained that helped to build my Multinomial Logit model.

Keywords: Statistical Model, Mobility, Politomy, Dummy, Multinomial Logit, Principal Component Analysis, R-Macfaden, Kaiser Meyer, Cox Shell, Wald's Coefficient.

CAPÍTULO 1

PROBLEMA DE INVESTIGACIÓN

1.1 Introducción

Indudablemente, el transporte público en el Ecuador se ha constituido en una parte básica para el desarrollo de las diferentes ciudades y para la descentralización de las actividades en las principales urbes del país. El transporte que se ha modernizado y se ha tecnificado, debe garantizar el traslado oportuno de cada uno de sus usuarios desde el punto de partida hacia su lugar de origen.

Las entidades encargadas de garantizar una movilidad segura y adecuada en Ecuador, son: la Agencia Nacional de Transito y Gobiernos Municipales a través de sus direcciones de Movilidad quienes, a través de reglamentos, políticas, ordenanzas, etc. tratan de mejorar las condiciones de movilidad humana, con el propósito de tener sistemas de transporte publico eficiente, accesible y económico. Sin embargo, no se puede determinar a criterio público que este sea el óptimo.

La estrategia que se empleó para el presente estudio es tratar de analizar tres aspectos dentro de lo que es Transporte público de pasajeros como son: Infraestructura, calidad del servicio y ascenso y descenso de pasajeros.

La Investigación se realizó tres fases, la primera se enfocó en realizar: El problema de Investigación, Justificación del problema a investigar, luego se realizó una amplia búsqueda bibliográfica que se constituyó en los antecedentes investigativos; una segunda fase se constituyó en la elaboración del Marco Metodológico y Obtención de Resultados; y finalmente se terminó con conclusiones y recomendaciones.

La principal problemática radica en la poca información bibliográfica que existe en relación a la aplicación de modelos estadísticos que contribuyan a dar solución a un aspecto muy importante dentro de la sociedad y movilidad como es el transporte público de pasajeros.

1.2 Justificación

En la ciudad de Riobamba existe un problema en brindar un eficiente servicio de transporte público se realizó un modelo un modelo estadístico que permita valorar este servicio.

Los resultados de la investigación ayudaran a mejorar varias aristas que se relacionan directa e indirectamente con el transporte público: Al mejorar la movilidad en el cantón Riobamba implicara que su comercio se dinamice adecuadamente obteniendo una mejor economía para el cantón.

El estudio además ayudara a que se busquen nuevas rutas que cubran de mejor manera toda la ciudad lo que implicara un mejor desarrollo urbano y social, tomando en cuenta que el principal objetivo del servicio de transporte público es toda la sociedad de un lugar geográfico.

La investigación también ayudara a que a futuro se piense en el uso de transporte alternativo debido a un sustancial crecimiento poblacional y que este tipo de transporte debe ser amigable con el medio ambiente.

Los beneficiarios de los resultados alcanzados serán directamente todos los ciudadanos del cantón Riobamba, a razón de que, si aplica mejoras al transporte público de pasajeros, se va a ver reflejado en que los usuarios se van a sentir cómodos, seguros e incluso identificados por utilizar este medio de movilidad.

1.3 Objetivos

1.3.1 General

- Modelizar cualitativa y cuantitativamente las variables de incidencia del transporte público mediante análisis multivariado para identificar líneas adecuadas del transporte público en la ciudad de Riobamba.

1.3.2 Específicos

- Determinar las variables que mejor incidan en el modelado estadístico
- Aplicar análisis de componentes principales a un conjunto de variables endógenas y exógenas
- Aplicar análisis factorial al conjunto de variables endógenas y exógenas
- Elaborar una línea base con la información primaria y secundaria del transporte público en la ciudad de Riobamba

CAPÍTULO 2

ANTECEDENTES INVESTIGATIVOS

El estudio de la movilidad es de gran interés de los investigadores en distintas áreas, esto a causa de la espacialidad de los resultados de cada persona, motivan a conocer las diferentes causas o factores que intervienen en la salida de una persona, pues cada uno tiene un lugar específico y la ruta que desea seguir hasta llegar a punto de destino (Orellana, 2016).

En las diferentes urbes todo individuo sale de su hogar con el propósito de realizar diferentes actividades como: ir al trabajo, realizar compras, visitar amigos, etc. Esto genera que se produzca el efecto de estudiar la movilidad por parte de investigadores y en base a esto tomar decisiones adecuadas y oportunas respecto a los modos de movilizarse como son: vehículos propios, transporte urbano, a pie y el uso de la bicicleta.

La movilidad en zonas urbanas es necesario y queda descrita de tal manera que identifica claramente su influencia en la calidad de vida, pues se desplaza personas y bienes, esta demanda crece en ciudades que están en vías de desarrollo (Romero & Lugo, 2018).

En los centros urbanos, el hecho de existir mayor demografía y por ende un mayor comercio, implica a que se exija una adecuada movilidad para mejorar la calidad de vida de sus pobladores, para esto se debe tomar en cuenta que existe normas y ordenanzas que ayudan a cumplir con la movilidad que todos deseamos.

El análisis y evaluación de la calidad del servicio de transporte urbano es compleja pero necesaria, en la práctica emplean valoraciones que se basan en la medición de factores de acuerdo al desempeño del sistema de transporte, teniendo en cuenta que las personas o usuarios son los actores principales (Sánchez & Romero, 2010).

Las personas o usuarios del transporte urbano son el punto central o el elemento básico en el estudio de la calidad de este servicio, pues son los principales actores que

interactúan día a día con este medio, sin embargo, es necesario investigar los factores principales que determinen la eficiencia y calidad de cada uno.

El acelerado crecimiento poblacional genera necesidades básicas a ser cubiertas por quienes dirigen los pueblos, los sistemas de urbanización que se originan en función de los requerimientos de los habitantes de una ciudad trae como consecuencia la necesidad de movilidad (José & Gustavo, 2019).

Este crecimiento exponencial que actualmente experimentan las ciudades de cualquier país, genera que cada cierto tiempo se realice un estudio de los recorridos que realiza cada línea de transporte público, porque el crecimiento exponencial de la población va a generar nuevas demandas hacia el transporte.

Una de las opciones más comunes para este tipo de estudio son los modelos matemáticos, estas dependerán del tipo de variable ya sea cuantitativa, cualitativa o ambas, pues generar un modelo que represente de manera óptima, la variable en estudio (Padilla, Llamuca, Calderón, & Villamarin, 2019).

La calidad del servicio puede ser interpretada mediante un modelo de regresión ya sea con respuesta numérica o categórica, sin embargo, no todas estas variables aportan de manera significativa al modelo, estas dependerán del encuestado, esto es porque cada uno tendrá una perspectiva diferente de cada una de estas.

Según Padilla, Llamuca, Calderón y Villamarín (2019) realizan un estudio de la valoración del transporte urbano en la ciudad de Riobamba, concluyendo la deficiencia del servicio, donde concluyen que la calidad del servicio es pésima, no cubren todas las zonas urbanas, tiempo de espera de más de 10 minutos, además, recomiendan dar capacitaciones acerca del buen trato y la convivencia entre choferes y usuarios.

El análisis de regresión múltiple se enfoca para variables cualitativas y cuantitativas, esto para facilitar el ajuste del modelo y sus respectivas predicciones, los modelos regresión logística binaria y multinomial trabajan con variables de respuesta categoría de dos o más categorías respectivamente, lo que nos permitiría valorar de forma efectiva el transporte público de pasajeros.

La caracterización de la valoración del transporte urbano es de respuesta categórica Politémica, con el fin de interpretar de manera óptima los resultados obtenidos se cuantifica la calidad de ajuste mediante el pseudo R^2 de Mac-Fadden, medida semejante al coeficiente de determinación utilizada para variables cuantitativas y de la misma manera representa la calidad del ajuste obtenido con el modelo ajustado (Pando & San Martín, 2004).

La percepción de los usuarios ante el servicio de transporte urbano es muy importante, pues la convivencia entre pasajeros puede verse afectado por distintos factores y una de ellas es el acoso a las mujeres, agresiones, robos, aglomeración y entre otros, cada una de estas perjudican la valoración del transporte urbano es por eso que para determinar la eficacia del modelo se utiliza el R^2 de Mac-Fadden.

Buscar las posibles soluciones ante problemas causados por el chofer y pasajeros o entre pasajeros es una de las prioridades en distintas ciudades, uno de los ejemplos es la me México donde implementan un transporte solo para mujeres, y la percepción de las féminas es de carácter principal pues comentan que no se cumple con las disposiciones y que aún pueden verse hombres en los asientos lo cual las trae inseguridad y a la vez molestia (Romero, Sánchez, & Lara, 2016).

Uno de los objetivos del transporte urbano debe ser la seguridad y el bienestar de los usuarios pues de esta manera brindan un servicio excelente y de calidad, ya que esta valoración viene dada por la satisfacción de los usuarios (Llanos, Llanos, & Molina, 2013).

Llegar a la satisfacción entre oferta y demanda se logrará mediante la convivencia sana entre estos dos grandes grupos, otra de las partes más importantes es la adecuación del transporte, para que los usuarios viajen cómodos a su destino, si esto no se cumple entonces el pasajero estará insatisfecho.

Por naturaleza el ser humano necesita transportarse de un lugar a otro, esto resulta complicado en ciudades extensa y con recorridos de larga distancia, esto dependiendo

de las actividades que desee realizar la persona, para ello busca optimizar su economía y tiempo (Delgado, Prado, & Ochoa, s.f).

Mientras un pasajero pase más tiempo dentro del servicio, su percepción ante esta sería más óptima, en esto el usuario puede analizar su entorno como: la limpieza del bus, asientos en buen estado, un sitio donde se deposite la basura y la aglomeración que pueda sufrir en horas pico, además el incumplimiento de horarios y rutas es una señal de insatisfacción.

El modelo Servqual es utilizado para medir la satisfacción de los usuarios frente al servicio de transporte, dicho método constituye en construir una medida subjetiva en el caso de expectativa y la percepción de los clientes, a lo cual se concluye que cada usuario tiende a calificar de distinta manera y de ello se concluye realizar estudios objetivos que ayuden a mejorar el pronóstico de esta información (Vega, Rivera, & Malaver, 2017) .

Una de las dificultades que se ha establecido con el modelo Servqual, es el trabajo con medidas subjetivas, esto es por el desarrollo de variables cualitativas para medir la calidad del servicio del transporte urbano, sin embargo, no se deja atrás la idea de implementar análisis que correlacione a las variables en estudio, sin embargo, un mejor estudio se podrá hacer mediante un modelo.

En la ciudad de Popayán realizan un análisis modal de elección discreta del transporte público y privado, donde utilizan el modelo Logit Multinomial donde las variables más significativas resultaron costos de transporte, el tiempo de dura en el servicio, el ingreso mensual del usuario y la edad de la persona, de estas el género y ser jefe de hogar no resultaron significativas (Fajardo & Gómez, 2015).

El modelo Logit Multinomial representa de manera óptima del servicio de transporte urbano, con variables de percepción del cliente sin importar su género ni la posición ocupada en el núcleo familiar, en este trabajo la variable ingreso mensual resulta significativo pues el estudio se enfoca a empresas públicas y privadas que brindan el servicio de movilidad, esto se logra cuando se realiza un modelo generalizado.

En el campus de la Universidad Francisco de Paula Santander Ocaña ubicada en Colombia realizan un estudio para establecer un nuevo modo de movilización para ello realizan ajustes a las variables obtenidas a modelo Logit Multinomial y Logit anidado, esto fue a que la mayoría de la población en estudio dio una negativa ente los transportes ya establecidos, allegando de mejor manera el nuevo modo de trasporte que afirma eliminar las incomodidades sentidas con otros (Duarte, Mendoza, Arévalo, & Guerrero, 2015).

Una de las preocupaciones de todos los usuarios es tener que esperar al bus que no cumple con los horarios establecidos, esto causa la aglomeración en las paradas y dentro de las unidades, lo cual genera incomodidad e inseguridad.

En la ciudad de Buenos Aires se realiza un trabajo donde pretende estudiar la reacción de los pasajeros de transporte urbano al modificar algunas variables como tiempo de viaje, costos, ingresos, etc. De esta manera evalúa el impacto de la demanda en los distintos modos de transporte, sin embargo, los resultados obtenidos no fueron favorables y de esto indica que solo serán de carácter indicativo pues no se pudo llegar a conclusiones más claras ante el problema estudiado (Garcia, 2005).

En la actualidad se nota el crecimiento del parque automotor a causa de la movilidad urbana, el trasporte se ve reducido, existe rebajas en los pasajes y autobuses que se enfrascan en la guerra del centavo, todas estas anomalías han hecho que el usuario tome otro tipo de alternativas, por lo cual es importante que se tome en cuenta resultados que son generados por modelos.

En Santa Marta, Colombia hacen un estudio de la satisfacción de los usuarios frente al sistema de transporte unificado, para ello se eligen las variables del modelo SERVQUAL, relacionando con el costo, servicio, estado de vehículo, entre otras, y como era de esperarse las personas no están conforme con el servicio recibido (Rodríguez , Chacón, & Orozco, 2018).

Uno de los grandes problemas que se vive en la actualidad es la falta de respeto mutuo, esto ocasiona el descontrol en trasportes urbano lo que implica una mala calificación

a su servicio, sin embargo, las pocas personas que sienten satisfacción es un pilar importante para futuros cambios en el comportamiento humano.

La percepción de los usuarios ante el servicio de transporte urbano y suburbano en la ciudad de Madrid esto se realiza con el fin de identificar las razones de preferencias de los pasajeros en elegir un medio de transporte. Se determina que existe una inclinación por el metro, sin embargo, tienden a sugerir mejoras en el entorno del transporte (limpieza, ambiente, etc.) (Delgado, Gómez, & De Esteban, 2019).

Es muy importante tener en cuenta los servicios del transporte que brinda una empresa o compañía en una ciudad pues la preferencia tendrá más significado, mientras haya competencia, además se debe tener en cuenta como mejora sustancialmente el entorno del transporte público, pues es una de las causas que insatisface a los usuarios.

Identificar de manera óptima las variables que pueden explicar de manera significativa es lo primordial en cualquier trabajo de investigación, para ello se establecen directrices que ayuden a obtener las variables recomendadas al momento de realizar la encuesta, y las respuestas de los usuarios implicara mucho la significancia de estas características (Brahmaiah, Prasad, & Srinivas, 2017).

Sin duda, determinar las variables que ayuden a identificar un correcto funcionamiento en mi modelo es fundamental a hora de establecer correspondencias y a su vez esto sirva para un correcto funcionamiento en mi modelo.

La existencia de aglomeraciones es un problema a resolver y esto es la principal preocupación de los usuarios de la población urbana en Noruega en la región de Oslo, los resultados arrojan que la probabilidad de experimentar robos en los modos de transporte colectivo es más alta que aquellos que preferían los modos privados, además el miedo de sufrir accidentes y actos de terrorismo, sin embargo la preferencia del transporte no se relaciona con las preocupaciones y riesgos dentro del bus público o privado (Roche, Rundmo, Foss, & Moe, 2013).

En diferentes urbes del mundo, la preocupación por mejorar el transporte público es un problema a resolver, existen características que deben someterse a un estudio profundo como análisis correlacionales o análisis predictivos.

Existen ocasiones donde la mayoría de los usuarios del transporte urbano son mujeres, y se dice que las mujeres son más sensibles a los problemas que presentan estos servicios como el tiempo que dura el servicio y la falta de atención al horario establecido, estas variables indican que mientras esperan dependen de los teléfonos móviles con su auriculares, esto determina que se debe tener en cuenta las necesidades de los hombres y mujeres (Chowdhury & Van Wee, 2020).

En ocasiones no es muy importante el realizar un estudio por género, en el caso del transporte público es fundamental determinar características y tendencias de problemas que vayan asociadas al género, pues la gran mayoría de las personas que utilizan este medio de transporte son las mujeres.

En Gdynia realizan un estudio tomando como muestra el 1% de sus habitantes comprendidos en edades de dieciséis a setenta y cinco años, quienes evaluaron al servicio del tren de manera positiva y el de transporte urbano de manera regular, esto a causa del tiempo que dura el transporte ya que la ciudad ha estado en crecimiento (Hebel & Wolek, 2016).

El target que se toma en cuenta para un estudio nos permitirá tener evidencias más cercanas a la realidad, en nuestro caso según estudios preliminares y por recomendación de expertos se expuesto tomar un rango de dieciséis a setenta y cinco años para obtener información.

El análisis de indicadores de transporte urbano toma relevancia en la ciudad de Dhaka donde argumentan que para alcanzar una ciudad sostenible el transporte debe ser de calidad y sostenible pues equilibra lo social, económico y ambiental hacia una meta. Identifico un conjunto de 14 indicadores, de los cuales el costo y el tiempo son los más relevantes (Munira & San Santoso, 2017).

Los indicadores como se conoce son medidas que nos permita evaluar de forma efectiva una investigación, en el caso del transporte público de pasajeros existe un sinnúmero de variables, las cuales pueden o no incidir en el modelo que se pretende construir, por eso es importante realizar una caracterización de las variables más importantes que incidan en mi modelo.

CAPÍTULO 3

MARCO METODOLÓGICO

3.1 Ubicación

La ciudad de Riobamba está ubicada en la Cordillera de los Andes, centro geográfico del Ecuador, a una altura de 2750 msnm centro de la hoya del Chambo, esta se encuentra rodeada por majestuosos volcanes como el Chimborazo, Carihuairazo, Tungurahua y el Altar.

Según el censo poblacional INEC (2010) Riobamba cuenta con 146 324 habitantes, de los cuales, 107 114 hab. son considerados para el estudio.

El área de estudio corresponde a la parte urbana de la ciudad de Riobamba por donde transitan las distintas líneas de transporte, que brindan el servicio de a la ciudadanía.

La ciudad de Riobamba cuenta con 5 parroquias urbanas: Veloz, Lizarzaburu, Maldonado, Juan de Velasco y Yaruquies, el servicio de transporte urbano realiza recorridos por distintas rutas con el fin de brindar servicio a toda la población que utilice estos sistemas de circulación.

Actualmente la parte urbe cuenta con 16 líneas de transporte, y son los encargados de transportar a la población riobambeña a los distintos puntos de la urbe. Cada una de estas tiene distintos sitios de salida y llegada y día a día planean mejorar el servicio que ofrecen.

3.2 Equipos y materiales

La encuesta fue el material principal para la recolección de información utilizando el instrumento denominado cuestionario, y se utilizó una computadora para poder digitalizar toda la información recabada y realizar el análisis correspondiente en el software estadístico Rstudio con sus librerías *Hmisc*, *ggplot2*, *MVN*, *dplyr*, *DMwR2*, *openxlsx*, *MASS*, *nnet*, *tidyr*, *questionr*, *GPArotation*, *vegan*, *PerformanceAnalytics*, *centralImputation*, *readxl*.

3.3 Nivel o tipo de investigación

El tipo de Investigación utilizado para el presente estudio es de tipo exploratorio, basado en un enfoque cuantitativo y cualitativo.

3.3.1 Investigación Exploratoria

El trabajo de investigación será de nivel exploratorio porque la calificación del transporte público urbano es poco estudiada, además se desarrollará planificaciones para estudios futuros mediante la recolección de información más detallada que ayuden a definir de manera efectiva el objetivo de estudio.

3.3.2 Investigación Descriptiva

La investigación será descriptiva porque se analizará de forma detallada la percepción de los usuarios del transporte público, se fundamentará principalmente en el análisis de variables categóricas.

3.3.3 Investigación Explicativa

El trabajo de investigación es explicativo, pues se desarrollará un modelo que explique la valoración del transporte público por medio de una muestra representativa de los usuarios, además servirá de base, para estudios a futuro.

3.3.4 Investigación Correlacional

La investigación será correlacional porque busca medir el grado de relación entre la calidad del servicio de transporte urbano con las variables explicativas recolectadas,

3.4 Prueba de Hipótesis

Hipótesis nula: El modelo es significativo

Hipótesis alternativa: El modelo no es significativo

3.5 Población o muestra

La población en estudio son los usuarios de las distintas localidades de la ciudad de Riobamba quienes utilizan las líneas de transporte público, y son quienes darán contestación a las preguntas formuladas en el cuestionario.

La muestra será tomada de acuerdo al número de habitantes, en las 5 parroquias urbanas de la ciudad de Riobamba, la población considerada para la investigación son ciudadanos que tengan edades de 15 años en adelante.

La muestra se calculó de acuerdo a la siguiente ecuación de cálculo.

$$n = \frac{NZ_{\alpha}^2 * p * q}{d^2 * (N - 1) + Z_{\alpha}^2 * p * q}$$

dónde:

N = tamaño de la población

Z = nivel de confianza,

P = probabilidad de éxito, o proporción esperada

Q = probabilidad de fracaso

D = precisión (Error máximo admisible en términos de proporción).

Según el censo poblacional del 2010 Riobamba cuenta con un total de 107114 habitantes quienes conforman el grupo de edad de los 15 años en adelante.

$$n = \frac{(107114) * (1,96^2) * 0,5 * 0,5}{4,924E - 02^2 * (107114 - 1) + (1,96^2) * 0,5 * 0,5} = 395 \text{ habitantes}$$

La encuesta se realizó durante el día en horarios de mayor circulación peatonal.

El transporte urbano cubre las 5 parroquias urbanas de la ciudad de Riobamba, y para el levantamiento de información se dividió los $\frac{395 \text{ hab.}}{5 \text{ parroquias}} = 79$, de esta forma se obtiene un total de 79 encuestas por cada parroquia.

Los datos de la segunda matriz fueron tomados en referencia en cada uno de los tramos (parada a parada), respecto a los pasajeros que subieron, bajaron y no accedieron a utilizar el servicio; además se recabo el estado de las vías y paradas. De esta manera se tuvo una matriz de 573 datos recolectados con 14 variables.

3.6 Recolección de Información

La recolección de información se realiza por medio de la técnica de encuesta y utilizando el instrumento cuestionario el mismo que contendrá preguntas cerradas para facilitar la eficacia y eficiencia del cuestionario en la obtención de la información, dado que la población de estudio es de movimiento constante.

Además, se recolecto la información de las diferentes líneas de transporte en base a cada uno de sus tramos (entre paradas), esto con el fin de identificar el comportamiento peatonal ante la llegada de un bus a la parada correspondiente.

La recolección de la información se fundamentó en tres aspectos importantes que analiza el transporte como es: Calidad del servicio, Infraestructura, y ascenso y descenso de pasajeros.

3.7 Procesamiento de la información y análisis estadístico

EL objetivo principal del estudio es modelizar cualitativa y cuantitativamente las variables que inciden en el transporte público mediante análisis multivariado para valorar el transporte público en la ciudad de Riobamba. De acuerdo a la variable respuesta que es Politémica, se consideró en analizar el modelo Logit Multinomial, además se consideró las pruebas de hipótesis de dependencia entre las variables con el

fin de verificar la dependencia de las variables explicativas de acuerdo a la variable respuesta.

Para poder realizar el análisis el modelo Logit Multinomial se trabajó con variable dummy de acuerdo a cada una de las categorías de cada variable con respuestas de 1 y 0, es decir que 1 quiere decir que se cumple la condición de la variable y 0 quiere decir que no se cumple la condición de la variable, de esta manera se ajusta el modelo con todas las variables ficticias o dummy y se realiza las pruebas de los coeficientes de Wald, mediante el cálculo del valor estadístico p con una prueba de significancia de 0.05.

Además, se consideró la validación del modelo utilizando el Pseudo R² de MacFadden, Cox-shell, Nagelkerke y la tabla de clasificación correcta e incorrecta.

Para el análisis de la matriz 2, que contiene información recolectada en cada uno de los tramos de las diferentes líneas de transporte, se realizó el análisis descriptivo según el tipo de variable, además se verifico el supuesto de normalidad univariada y multivariada de las variables numéricas detallando la prueba de esfericidad de Bartlett y Kaiser-Meyer-Olkin (KMO) las cuales son estadísticos importantes para la aplicación y se procedió a realizar el análisis Factorial y Componentes principales.

3.8 Variables respuesta o resultados alcanzados

Para cumplir los objetivos planteados se toma como variable dependiente la valoración del transporte urbano frente a las variables independientes: Inconvenientes de aglomeración, trato por conductor, kit de seguridad, tiempo de origen a destino y porque utiliza el transporte urbano; como resultado se desea obtener variables significativas que representen de manera óptima a la variable respuesta y pueda caracterizar de forma idónea la valoración del transporte público en la ciudad de Riobamba.

CAPÍTULO 4

RESULTADOS Y DISCUSIÓN

4.1 Análisis e interpretación

4.1.1 Imputación de datos faltantes

Una vez realizadas la encuesta se procede a realizar la tabulación respectiva de toda la información en la hoja electrónica Excel, seguidamente se importa al software estadístico Rstudio donde se analiza toda la información extraída empezando por los datos faltantes o datos vacíos.

```
colSums(is.na(calidad)) # Indica el número de NAs por cada variable
```

Tabla 4-1: Valores faltantes por variable

Número de datos faltantes por Variable	
VAL_TRANSPORTE	INCONVENIENTES_AGLOMERACION
0	1
TIEMPO_ORIG_DESTINO	TRATO_POR_CONDUCTOR
0	1
KIT_SEGURIDAD	PORQUE_UTI_t.u
2	2

Interpretación: En la tabla 4-1, se observa que la información recolectada no posee un gran número de faltantes, por tanto, se puede realizar una imputación de datos mediante la función *centralImputation* de R la cual reemplaza la información faltante según el tipo de variable: si es cuantitativa reemplaza por la mediana y si es cualitativa se reemplaza por la moda, de la siguiente manera.

```
centralImputation(calidad)
```

Tabla 4-2: Valores imputados por variable

Número de datos faltantes por Variable	
VAL_TRANSPORTE	INCONVENIENTES_AGLOMERACION
0	0
TIEMPO_ORIG_DESTINO	TRATO_POR_CONDUCTOR

	0		0
KIT_SEGURIDAD		PORQUE_UTI_t.u	
	0		0

Interpretación: En la tabla 4-2 se muestra el conteo de datos faltantes por cada variable, luego de realizar la imputación no existe información vacía, esto significa que la base fue rellenada con existo, la función *centralImputation* uso la moda para cubrir los vacíos existentes pues todas son variables cualitativas que contienen observaciones realizadas por cada uno de los usuarios según el cuestionario que se planteó.

4.1.2 Análisis descriptivo

- **Valoración de transporte**

Tabla 4-3: Frecuencias de la variable Valoración de transporte

Categoría	BUENA	INSUFICIENTE	MALA
Frecuencia	172	113	110
Proporción	0.435	0.286	0.278

Interpretación: La tabla 4-3 indica las frecuencias correspondientes a la valoración de transporte, donde el 172 de los encuestados califican el servicio del transporte público como buena, el 113 califica como insuficiente y 110 califica como mala, esto equivale al 43,5%, 28,6% y 27,8% respectivamente. A pesar de contar con una mayor proporción de usuarios que están conformes con el servicio, no es prueba suficiente de que la ciudadanía también tienen el mismo pensamiento pues al juntarse las proporciones de insuficiente y mala, se tiene más del 50% que esta inconformes con el servicio de transporte.

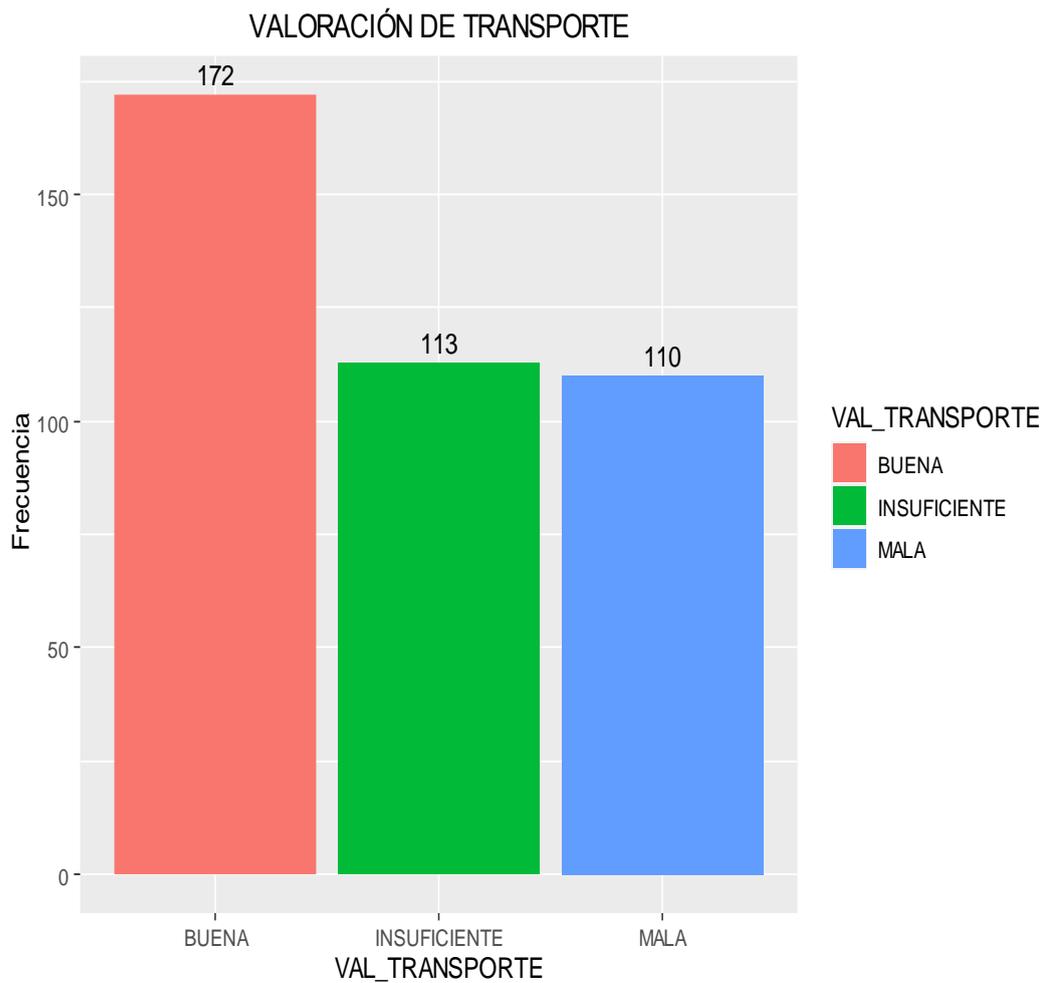


Figura 4-1: Frecuencia absoluta de la variable Valoración de transporte

El 43.5% de los encuestados califican el servicio del transporte público como buena, el 28.6% califica como insuficiente y 27.8% califica como mala, lo que representa a un total de 172, 113 y 110 personas respectivamente.

- **Inconvenientes de aglomeración**

Tabla 4-4: Frecuencias de la variable Inconvenientes de Aglomeración

Categoría	CASI NUNCA	CASI SIEMPRE	ESPORÁDICAMENTE	NUNCA	SIEMPRE
Frecuencia	67	19	82	217	10
Proporción	0.170	0.048	0.208	0.549	0.025

Interpretación: La tabla 4-4 indica las frecuencias y proporciones de inconvenientes de aglomeración que presenten los usuarios al momento de utilizar el servicio, según el análisis 217 personas respondieron que nunca hay aglomeración frente a 10 que tuvieron este problema, del mismo modo 67 dijeron casi nunca 19 casi siempre y 82 esporádicamente, esto indica que el transporte urbano en la ciudad opera con normalidad y sin mucha aglomeración.

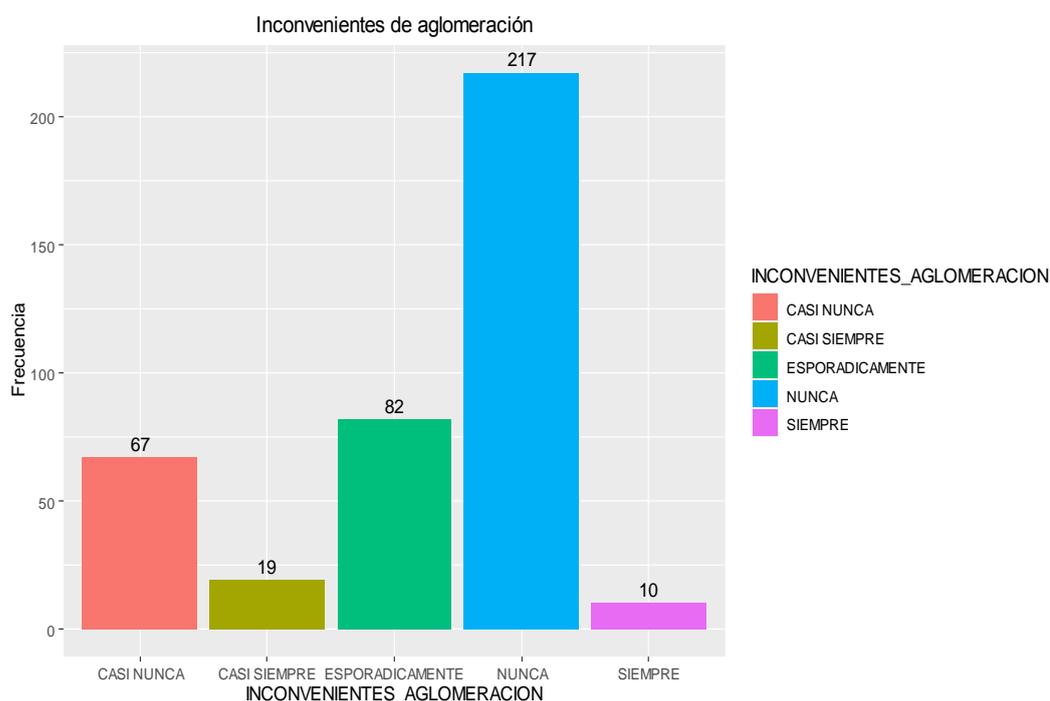


Figura 4-2: Frecuencia absoluta de la variable Inconvenientes de Aglomeración

El 17% de los encuestados mencionan que casi nunca existen problemas de aglomeración, el 4,8% indica que casi siempre, el 20,8% esporádicamente, el 54,9%

nunca y 2,5% indicaron que siempre, esto representa un total de 67, 19, 82, 217 y 10 personas respectivamente.

Tiempo de origen a destino

Tabla 4-5: Frecuencias de la variable Tiempo de origen a destino

Categoría	5 MIN	10 MIN	15 MIN	20 MIN O MAS
Frecuencia	16	56	96	227
Proporción	0.041	0.142	0.243	0.575

Interpretación: La tabla 4-5 muestra las frecuencias de los tiempos de viajes aproximados que realizan cada uno de los usuarios, de los cuales 16 personas utilizan por 5 minutos, 56 por 10 minutos, 96 por 15 minutos y 227 por 20 minutos o más. La mayor parte de los usuarios deben trasladarse con viajes que duran más de 20 según las necesidades y quehaceres de los pasajeros que obligan a transportarse distancias largas dentro de la urbe.

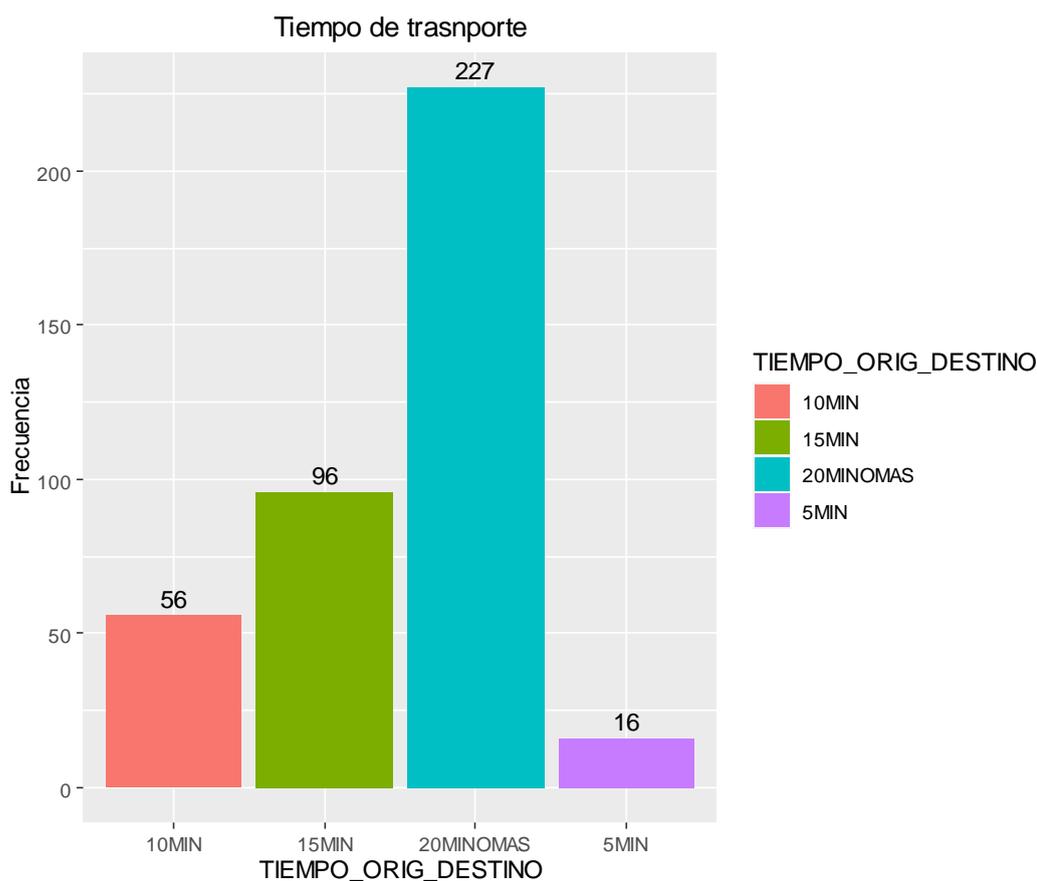


Figura 4-3: Frecuencias absoluta de la variable Tiempo de origen a destino

El 4.1% de los encuestados respondieron que el tiempo que dura el transporte es de 5 min, el 14.2% respondieron 10 min, el 24.3% dura 15 min y el 57,5% utiliza el transporte por más de 20 min, lo que es equivalente a un total de 56, 96, 227 y 16 pasajeros respectivamente.

Razón de utilizar el transporte público

Tabla 4-6: Frecuencias de la variable Razón de uso del transporte público

Categoría	BAJO COSTO	COMODIDAD	RAPIDEZ	SEGURIDAD
Frecuencia	328	24	27	16
Proporción	0.830	0.061	0.068	0.041

Interpretación: La tabla 4-6 muestra las frecuencias de la razón de uso del servicio de transporte urbano, donde 328 personas respondieron que utilizan transporte urbano

por bajo costo, 21 por la comodidad, 27 por rapidez y 16 por seguridad. El hecho de uso de transporte muestra la economía baja o que el pasajero piensa en su bolsillo y la de su hogar, por otra parte, se notó la inseguridad, lentitud e incomodidad que deben soportar día a día, personas que optan por este tipo de servicio.

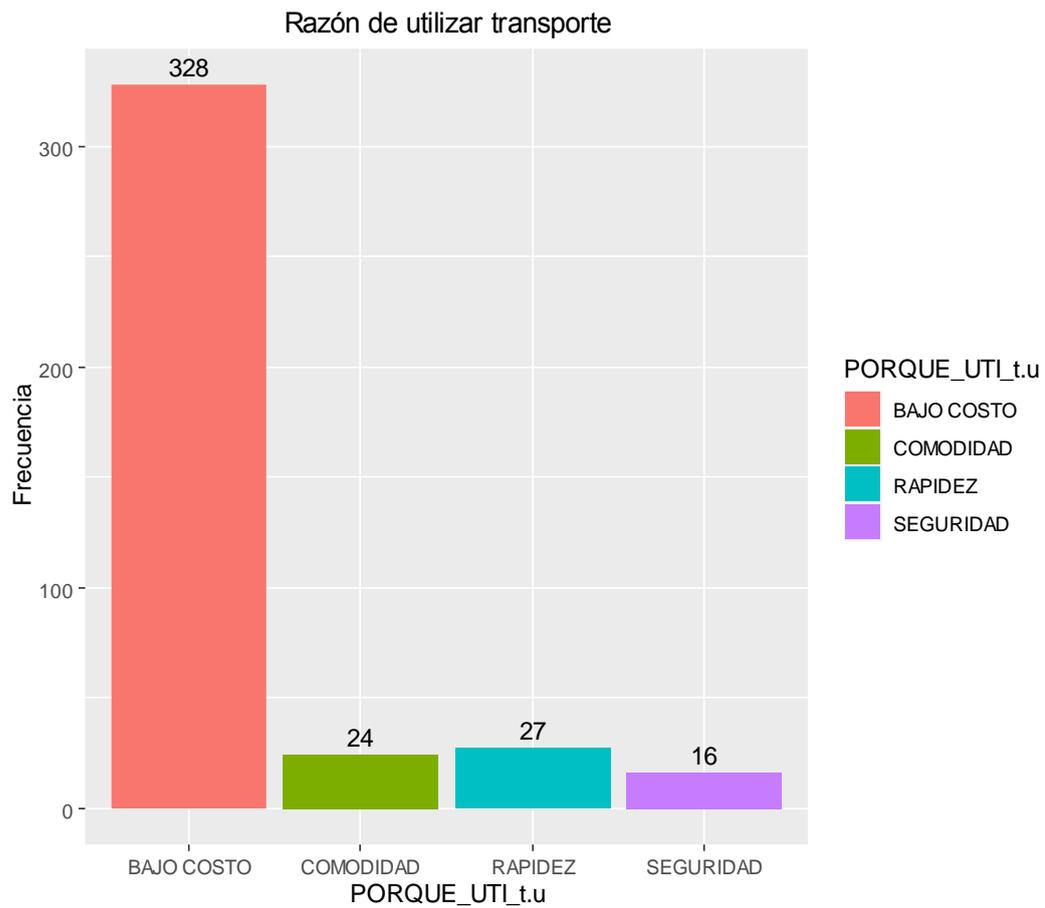


Figura 4-4: Frecuencias absoluta de la variable Razón de utilizar transporte

El 83% de los encuestados respondieron que utilizan el transporte público por bajo costo, el 6.1% por comodidad, el 6.8% por rapidez y el 4.1% respondieron que, por seguridad, esto equivale a un total de 328, 24, 27 y 16 personas respectivamente.

Trato por conductor

Tabla 4-7: Frecuencias de la variable Trato por conductor

Categoría	AGRESIVO	AMABLE	INDIFERENTE
-----------	----------	--------	-------------

Frecuencia	59	130	206
Proporción	0.149	0.329	0.522

Interpretación: La tabla 4-7 indican las frecuencias de los usuarios frente al trato recibido de por el conductor, de los cuales 59 recibieron un trato agresivo, 139 tratados con amabilidad y 206 de manera indiferente. Los conductores tienden a tener un mal comportamiento con los usuarios del transporte urbano, pues la mayoría de personas reciben tratos indiferentes o agresivos, cuando deben tener respeto de manera mutua entre pasajeros y conductores.

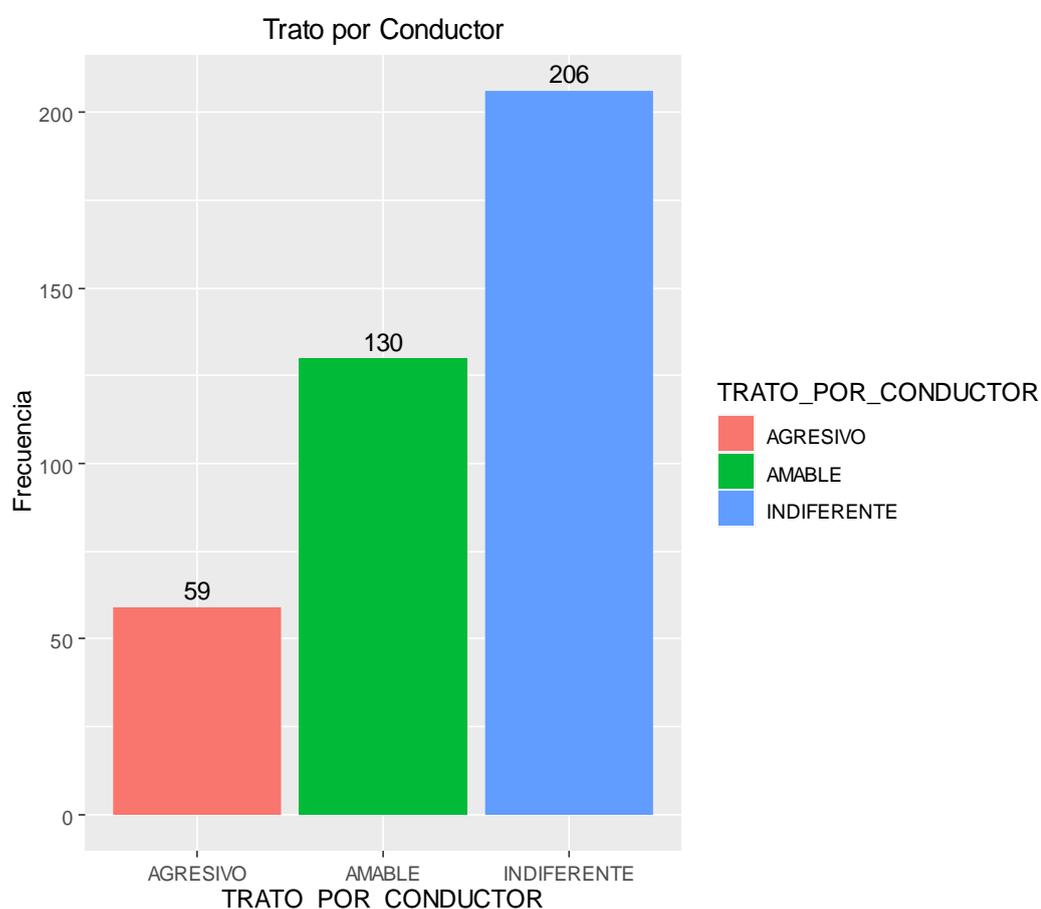


Figura 4-5: Frecuencias absoluta de la variable Trato por conductor

El 14.9% de los encuestados respondieron que el trato del conductor fue agresivo, el 32.9 % dijeron que fue amable y el 52.2% tuvo un trato indiferente, lo que equivale a un total de 59, 130 y 206 personas respectivamente.

Kit de Seguridad

Tabla 4-8: Frecuencias de la variable Kit de seguridad

Categoría	NO	SI
Frecuencia	224	171
Proporción	0.567	0.433

Interpretación: La seguridad en el transporte urbano es primordial para los usuarios del día a día, y la presencia de un kit de seguridad no es la excepción, pues cada unidad de transporte debe contar con ello, los usuarios respondieron que las unidades de transporte cuentan con un kit de seguridad solamente 171 y el resto contestaron que no. Autoridades competentes deben exigir la presencia de un kit de seguridad en cada unidad, de esta forma tanto conductores y pasajeros estarán seguros ante cualquier percance.

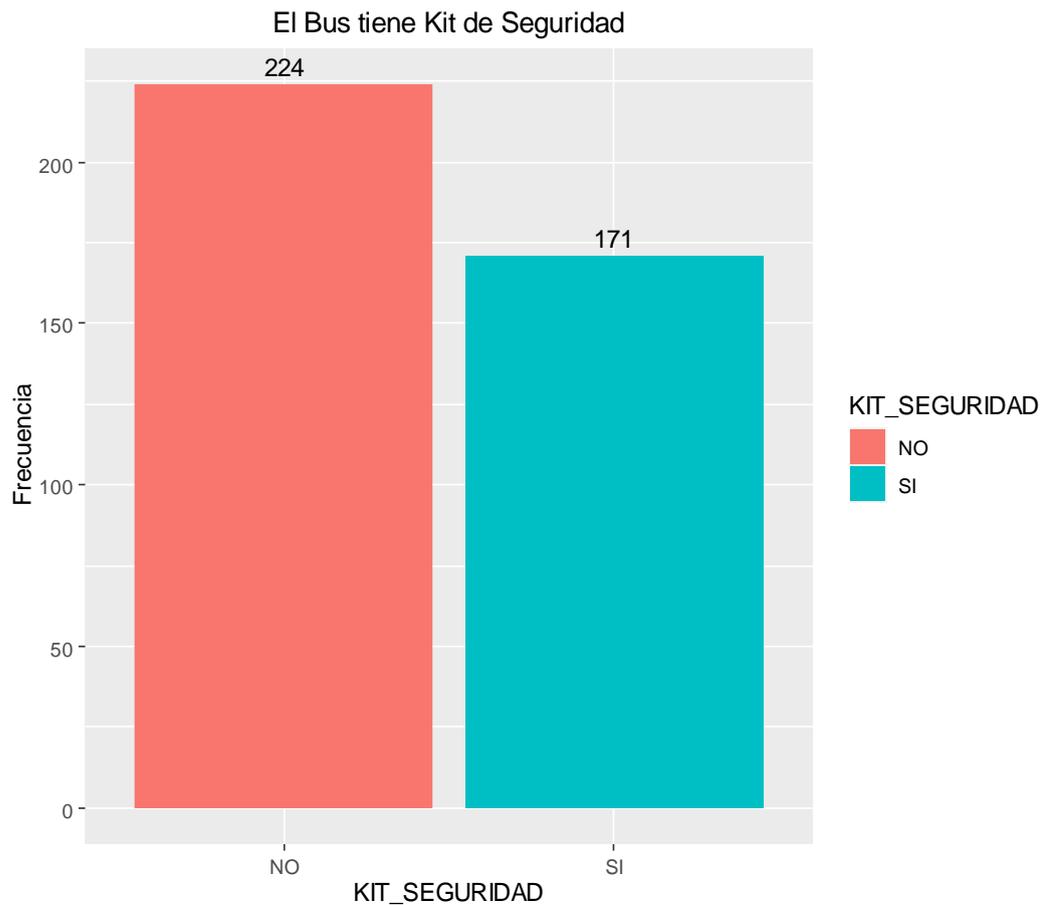


Figura 4-6: Frecuencias absoluta de la variable Kit de seguridad

El 56.7% de los encuestados respondieron que el bus no tenía kit de seguridad y el 43.3% de si tenía kit de seguridad, lo que equivale a un total de 224 y 171 personas respectivamente

4.1.3 Análisis descriptivo por pares de variables

Inconvenientes de aglomeración según la valoración de transporte

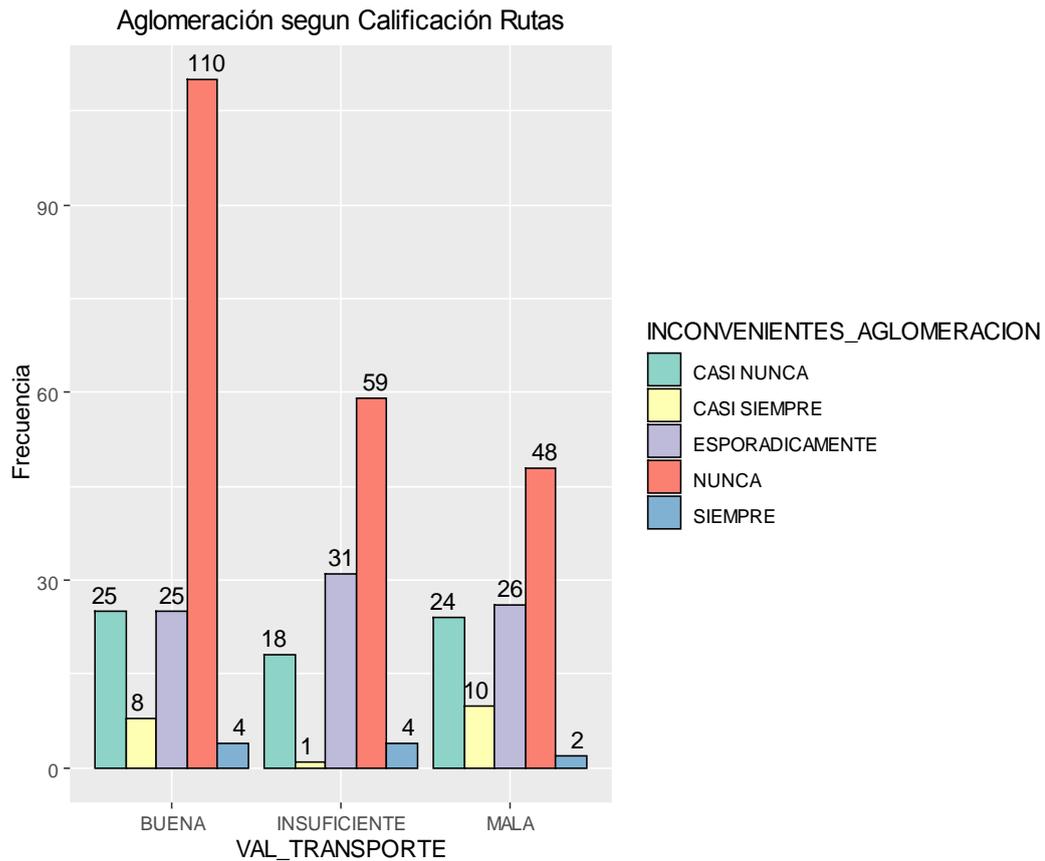


Figura 4-7: Frecuencias cruzadas de Valoración de transporte e inconvenientes de aglomeración

Los encuestados 110 calificaron como buena cuando la frecuencia de aglomeración no se presenta, cuando su frecuencia se ve reducido la valoración del transporte cambia a insuficiente y mala; la categoría casi nunca y esporádicamente no tienen una gran diferencia entre sus frecuencias; 4 encuestados que presentan inconvenientes de aglomeración siempre califican como buena e insuficiente y solo 2 como mala.

Tiempo de transporte según la valoración de transporte

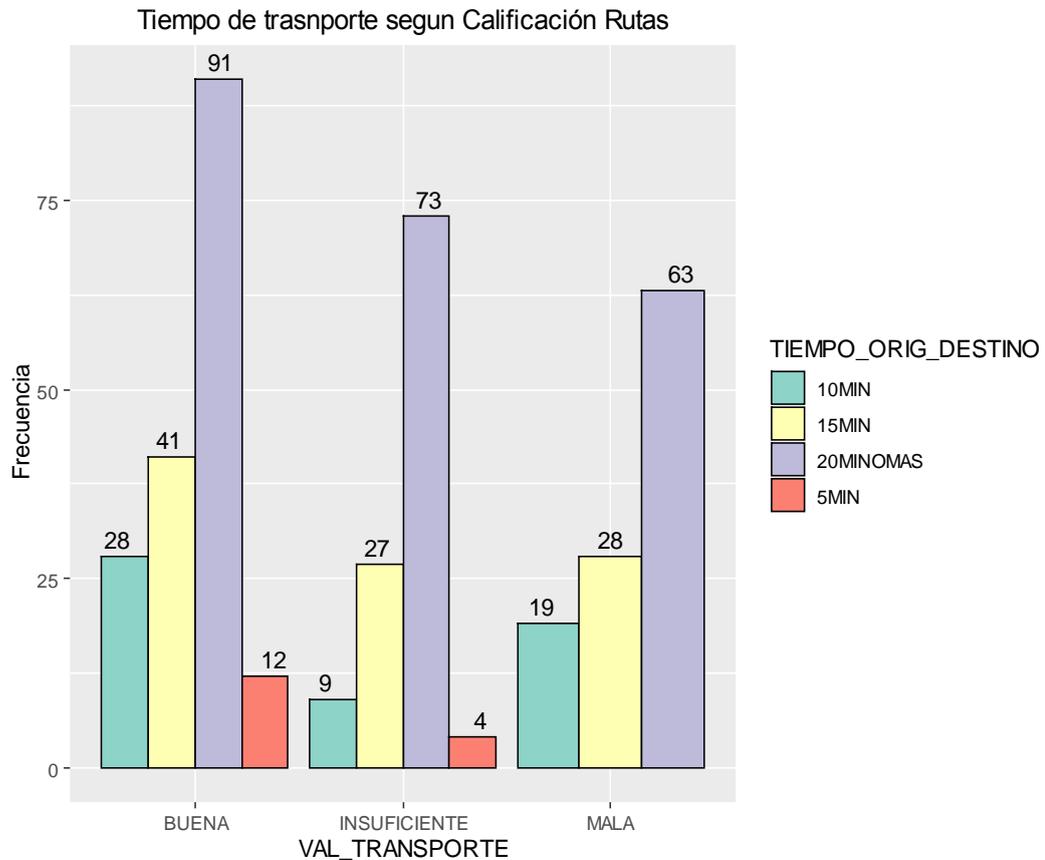


Figura 4-8: Frecuencias cruzadas de Valoración de transporte y tiempo de origen a destino

Si el usuario tarda 5 min en llegar a su destino valora como buena o insuficiente pero no mala, 91 pasajeros que tardan 20 min dentro del servicio dijeron que es bueno, 73 como insuficiente y 63 como mala, estas presentan las más altas en tras tres categorías, las personas que se transportan por 10 min y 15 min califican como buena, insuficiente en menor frecuencia, en este sentido mientras más tiempo pase en el servicio más acertada será su valoración.

Razón de utilizar transporte según valoración de transporte

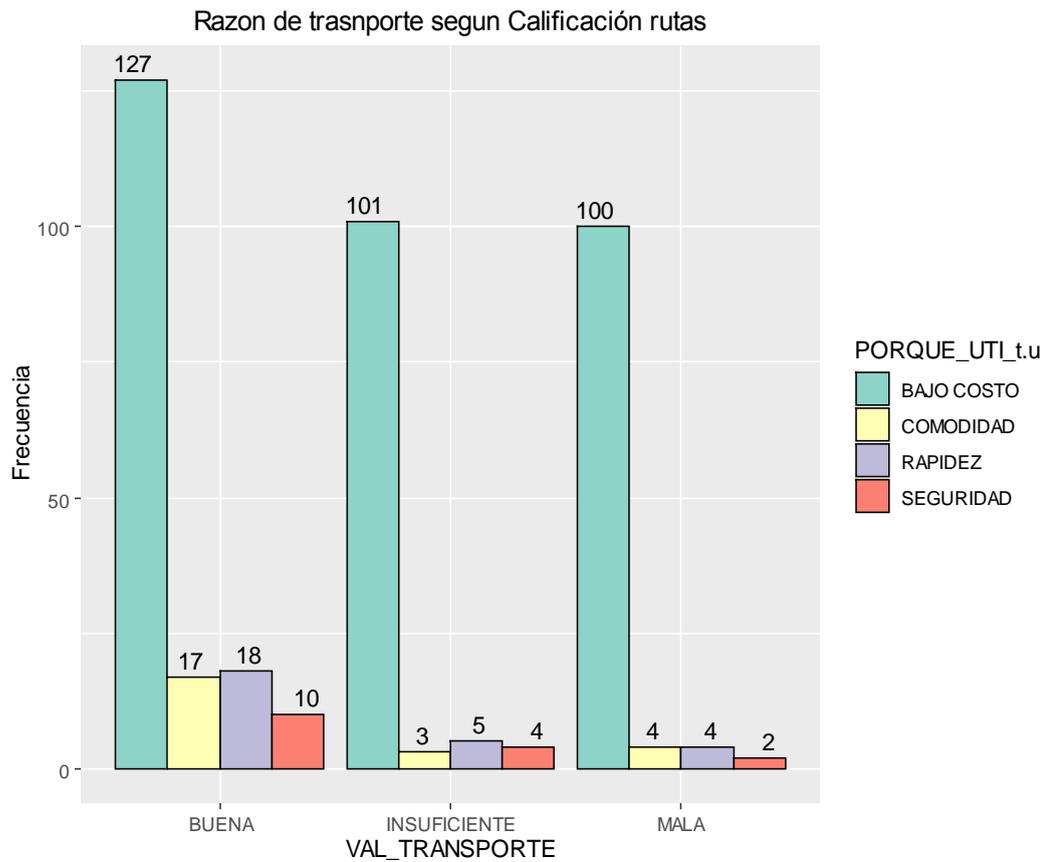


Figura 4-9: Frecuencias cruzadas de Valoración de transporte y Razón de uso del transporte

Sin importar que el transporte sea buena, insuficiente o mala las personas tienden a utilizar el transporte público por bajo costo con mayor frecuencia, caso contrario a la comodidad, rapidez y seguridad, esta última tiene una menor frecuencia en la calificación mala, esto puede ser porque no se siente seguro de viajar en transporte urbano.

Trato del conductor según la valoración de transporte

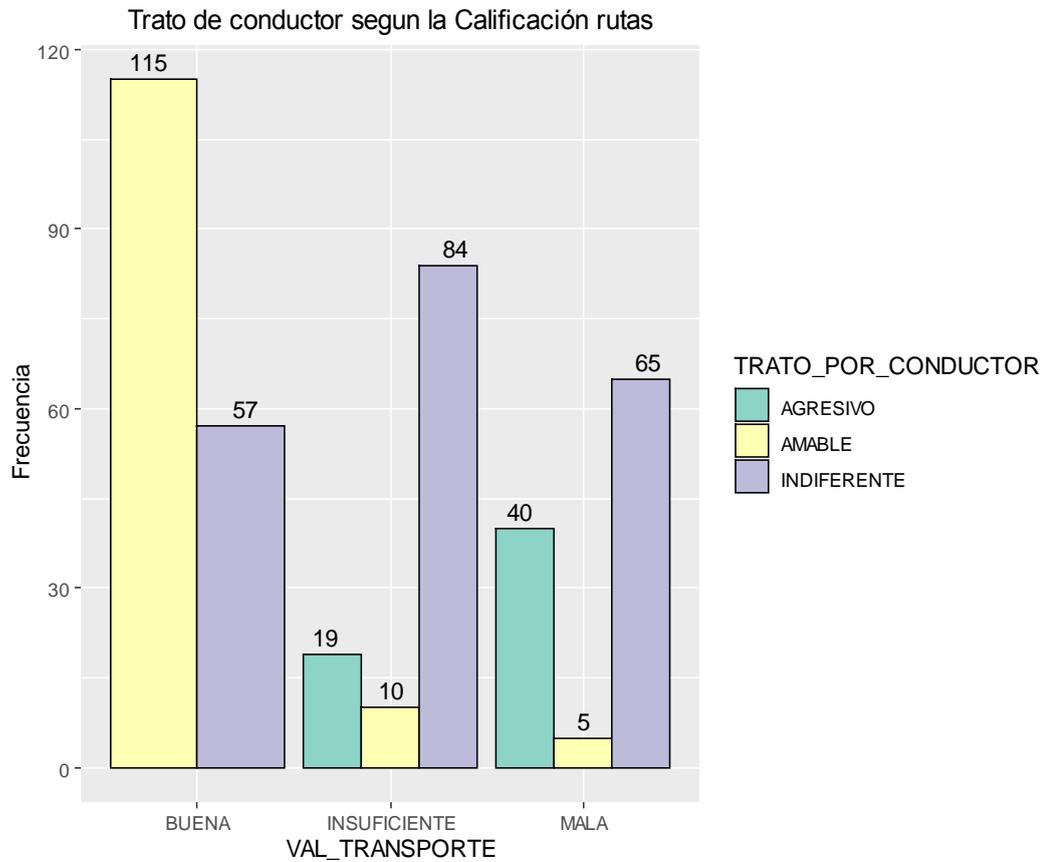


Figura 4-10: Frecuencias cruzadas de Valoración de transporte y trato por conductor

Según el grafico se observa que el trato del conductor incide en la calificación de las rutas; si el trato es agresivo con mayor frecuencia puede ser insuficiente o mala, caso contrario a que tenga una frecuencia menor de trato agresivo la valoración es buena, por otra parte, si el trato es amable tiene una valoración buena.

Kit de seguridad según la valoración de transporte

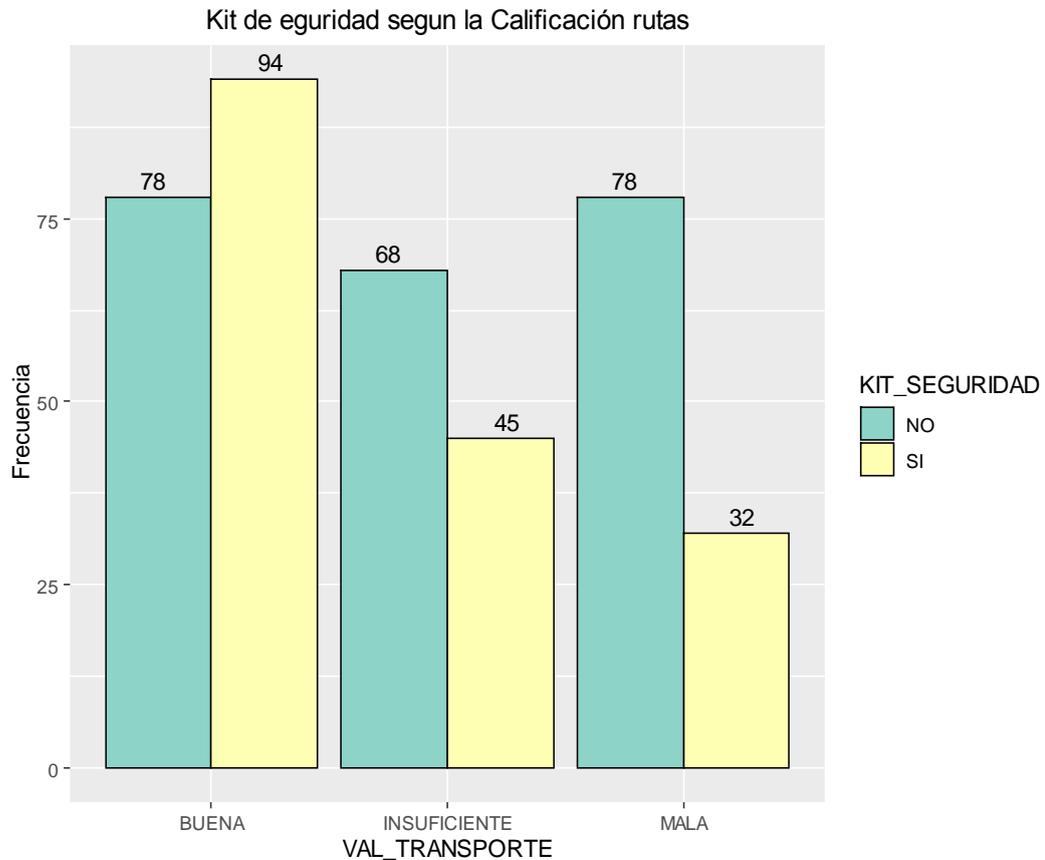


Figura 4-11: Frecuencias cruzadas de Valoración de transporte y Kit de seguridad

Las personas califican como buena cuando existe kit de seguridad y mala si no existe, mientras que califican como insuficiente y mala cuando no existe kit de seguridad dentro del transporte urbano.

4.1.4 Contraste de independencia Chi-cuadrado

Planteamiento de hipótesis

H_0 : Las variables son independientes

H_1 : Las variables son dependientes

Nivel de significancia de $\alpha = 0,05$

Estadístico de prueba

$$\chi_{obs}^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$
$$e_{ij} = \frac{\text{total filas} * \text{Total columna}}{\text{total}}$$

Decisión

Si valor $p < \alpha$ se rechaza H_0

Aplicación de la prueba en R

La siguiente función realiza la prueba de independencia Chi-cuadrado por medio de variables pareadas, la variable dependiente debe estar en la primera columna de la base de datos.

```
indep.test <- function(x){  
  p_val <- NULL  
  for (i in 1:ncol(x)) {  
    if(min(table(t(x[,1]), t(x[,i]))) < 5){  
      ft <- fisher.test((x[[1]]), (x[[i]]), simulate.p.value=TRUE, B=2  
e3)  
      p_val <- rbind(p_val, ft$p.value)  
    }  
    if(min(table(t(x[,1]), t(x[,i]))) >= 5){  
      ct <- chisq.test(x[[1]], x[[i]])  
      p_val <- rbind(p_val, ct$p.value)  
    }  
  }  
  library(dplyr)  
  return(as_tibble(p_val))  
}  
  
indep.test(calidad) %>% mutate("Pval < alpha" = V1 < 0.05) %>% muta  
te("Desicion" = if(V1<0.05) "Rechazo la hipótesis nula")
```

Tabla 4-9: Test Chi-cuadrado de independencia

Variable	P valor	$p < \alpha$	Decisión
VAL_TRANSPORTE	0.000500	TRUE	Rechazo la hipótesis nula
INCONVENIENTES_AGLOMERACION	0.00400	TRUE	Rechazo la hipótesis nula
TIEMPO_ORIG_DESTINO	0.0175	TRUE	Rechazo la hipótesis nula
TRATO_POR_CONDUCTOR	0.0005	TRUE	Rechazo la hipótesis nula
KIT_SEGURIDAD	0.0000901	TRUE	Rechazo la hipótesis nula
PORQUE_UTI_t.u	0.00350	TRUE	Rechazo la hipótesis nula

Interpretación: La tabla 4-9 indica la prueba de independencia y cabe recalcar que la primera fila representa la prueba de dependencia de la variable dependiente consigo misma, y las siguientes con las demás variables, en este caso se observa que para todas las variables se rechaza la hipótesis nula, por tanto, se concluye que las variable Calificación de rutas depende las demás variables: Inconvenientes de aglomeración, Trato por conductor, Tiempo de origen destino, Kit de seguridad, Razón de Utilizar el Transporte, por tanto se elabora un modelo empezando con cada una de estas variables. Si una de las variables no se rechaza la hipótesis nula significa que hay independencia entre las variables en estudio.

4.1.5 Análisis de modelo de predicción

Para buscar un modelo de predicción se la realiza de acuerdo al tipo de variable, en este caso la variable respuesta es de tipo cualitativo ordinal (valoración de transporte), esto porque la calificación del servicio se puede dar un orden jerárquico (mala, insuficiente, buena y muy buena)

De acuerdo a la bibliografía según el tipo de variable se puede aplicar un modelo Logit multinomial o modelo Logit Ordinal. Para el siguiente estudio se realizará el modelo Logit multinomial y ordinal.

Las variables explicativas serán transformadas a variables Dummy, esto será una nueva variable de acuerdo a cada categoría con respuestas de 1 y 0.

```

calidad_dummy <- calidad %>%
cbind(model.matrix(~INCONVENIENTES_AGLOMERACION -1, data =
calidad)) %>%
cbind(model.matrix(~TRATO_POR_CONDUCTOR -1, data = calidad))
%>%
mutate(KIT_SEGURIDAD = ifelse(KIT_SEGURIDAD == "SI", 1, 0)) %>%
cbind(model.matrix(~TIEMPO_ORIG_DESTINO -1, data = calidad))
%>%
cbind(model.matrix(~PORQUE_UTI_t.u -1, data = calidad))

```

Una vez transformada a variables dummy se tiene un total de 18 variables nuevas.

Tabla 4-10: Variables dummy según categorías de cada variable

Variables DUMMY	Valor
INCONVENIENTES_	num [1:395] 0 0 0 1 0 0 0 0 0 ...
AGLOMERACIONCASI NUNCA	
INCONVENIENTES_	num [1:395] 0 0 0 0 1 0 0 0 1 0 ...
AGLOMERACIONCASI SIEMPRE	
INCONVENIENTES_	num [1:395] 1 0 1 0 0 1 0 0 1 ...
AGLOMERACIONESPORADICAMENTE	
INCONVENIENTES_AGLOMERACIONNU NCA	num [1:395] 0 1 0 0 0 0 1 1 0 0 ...
INCONVENIENTES_	num [1:395] 0 0 0 0 0 0 0 0 0 ...
AGLOMERACIONSIEMPRE	
TIEMPO_ORIG_DESTINO10 MIN	num [1:395] 0 0 0 0 0 1 0 0 0 ...
TIEMPO_ORIG_DESTINO15 MIN	num [1:395] 0 0 0 0 1 0 0 1 0 0 ...
TIEMPO_ORIG_DESTINO20 MIN O MAS	num [1:395] 1 1 1 1 0 0 1 0 1 1 ...
TIEMPO_ORIG_DESTINO5 MIN	num [1:395] 0 0 0 0 0 0 0 0 0 ...
TRATO_POR_CONDUCTORAGRESIVO	num [1:395] 0 0 0 0 0 0 0 0 0 ...

TRATO_POR_CONDUCTORAMABLE	num [1:395] 0 1 0 1 0 0 0 0 0 0 ...
TRATO_POR_CONDUCTORINDIFERENTE	num [1:395] 1 0 1 0 1 1 1 1 1 1 ...
KIT_SEGURIDAD	num [1:395] 0 1 0 1 0 0 1 0 1 0 ...
PORQUE_UTI_t.uBAJO COSTO	num [1:395] 1 1 1 1 1 1 1 1 1 1 ...
PORQUE_UTI_t.uCOMODIDAD	num [1:395] 0 0 0 0 0 0 0 0 0 0 ...
PORQUE_UTI_t.uRAPIDEZ	num [1:395] 0 0 0 0 0 0 0 0 0 0 ...
PORQUE_UTI_t.uSEGURIDAD	num [1:395] 0 0 0 0 0 0 0 0 0 0 ...

Interpretación: La tabla 4-10 muestra la variable respecto a cada categoría, donde se crean nuevas variables que tiene como respuesta 1 y cero, realizado con el fin de analizar y ajustar con todas estas nuevas variables dummy un modelo predictivo que ayude a explicar la percepción de los usuarios ante el servicio de transporte urbano. El 1 y cero esto indica el cumplimiento o no se cumpla respectivamente.

Análisis de la significancia de las variables

```
res.mult <- multinom(VAL_TRANSPORTE ~ ., data = calidad_dummy)
```

Del resultado obtenido se analizan los estadísticos de Wald con el fin de calcular el valor p de cada una con el fin de analizar si las variables aportan o no al modelo.

Tabla 4-11: Estadísticos de Wald

Variable	INSUFICIE NTE	MALA
(Intercept)	7,59	-1,84
`INCONVENIENTES_AGLOMERACIONCASI SIEMPRE`	-1,46	0,61
INCONVENIENTES_AGLOMERACIONESPORA DICAMENTE	2,36	0,01
INCONVENIENTES_AGLOMERACIONNUNCA	2,76	-0,24
INCONVENIENTES_AGLOMERACIONSIEMPR E	0,84	-1,17
TRATO_POR_CONDUCTORAGRESIVO	84,19	85,00
TRATO_POR_CONDUCTORAMABLE	-24,16	-22,96
TRATO_POR_CONDUCTORINDIFERENTE	-19,57	-21,09

TIEMPO_ORIG_DESTINO10MIN	0,43	16,10
TIEMPO_ORIG_DESTINO15MIN	1,61	15,87
TIEMPO_ORIG_DESTINO20MINOMAS	3,38	21,14
TIEMPO_ORIG_DESTINO5MIN	0,46	- 8,81E+ 09
KIT_SEGURIDAD	-0,07	-1,53
`INCONVENIENTES_AGLOMERACIONCASI NUNCA`	1,90	0,69
`PORQUE_UTI_t.uBAJO COSTO`	3,16	2,24
PORQUE_UTI_t.uCOMODIDAD	0,52	0,78
PORQUE_UTI_t.uRAPIDEZ	-0,38	-1,71
PORQUE_UTI_t.uSEGURIDAD	1,05	-0,60

Interpretación: Una vez realizado el ajuste al modelo, se extrajo los estadísticos de Wald que se muestra en la tabla 4-11, dichos valores ayudan con el cálculo el valor p: si son negativos se calcula el p valor a cola izquierda y si es positivo se calcula a cola derecha. Esto con el fin analizar las variables que más influyen en la valoración del transporte urbano.

Tabla 4-12: Análisis de aporte al modelo de cada variable dummy

VARIABLES	Pval_1	des_1	Pval_2	des_2
`INCONVENIENTES_AGLOMERACIONCASI SIEMPRE`	7,20E- 02	N sig	2,71E- 01	N sig
INCONVENIENTES_AGLOMERACIONESPO RADICAMENTE	9,00E- 03	Sig	4,97E- 01	N sig
INCONVENIENTES_AGLOMERACIONNUNC A	2,90E- 03	Sig	4,05E- 01	N sig
INCONVENIENTES_AGLOMERACIONSIEMP RE	2,00E- 01	N sig	1,21E- 01	N sig
TRATO_POR_CONDUCTORAGRESIVO	0,00E+ 00	Sig	0,00E+ 00	Sig
TRATO_POR_CONDUCTORAMABLE	3,00E- 129	Sig	5,82E- 117	Sig
TRATO_POR_CONDUCTORINDIFERENTE	1,50E- 85	Sig	4,98E- 99	Sig
TIEMPO_ORIG_DESTINO10MIN	3,30E- 01	N sig	1,32E- 58	Sig
TIEMPO_ORIG_DESTINO15MIN	5,40E- 02	N sig	4,92E- 57	Sig
TIEMPO_ORIG_DESTINO20MINOMAS	3,60E- 04	Sig	1,65E- 99	Sig

TIEMPO_ORIG_DESTINO5MIN	3,20E-01	N sig	0,00E+00	Sig
KIT_SEGURIDAD	4,70E-01	N sig	6,27E-02	N sig
`INCONVENIENTES_AGLOMERACIONCASI NUNCA`	2,90E-02	Sig	2,45E-01	N sig
`PORQUE_UTI_t.uBAJO COSTO`	7,80E-04	Sig	1,26E-02	Sig
PORQUE_UTI_t.uCOMODIDAD	3,00E-01	N sig	2,17E-01	N sig
PORQUE_UTI_t.uRAPIDEZ	3,50E-01	N sig	4,33E-02	Sig
PORQUE_UTI_t.uSEGURIDAD	1,50E-01	N sig	2,74E-01	N sig

Interpretación: En la tabla 4-12 se analiza el aporte que tiene cada variable en el modelo de predicción, donde si el valor p supera al nivel de significancia no aporte al modelo caso contrario tiene un aporte primordial, si son significativos en ambas categorías se considera indispensable en el modelo.

De esta forma se tiene como primer modelo con las variables que son significativas en ambos casos y en uno de los casos, lo que detallamos a continuación:

$$\begin{aligned}
 VAL_RUTAS = & \\
 & INCONVENIENTES_AGLOMERACIONESPOREDICAMENTE + \\
 & INCONVENIENTES_AGLOMERACIONNUNCA + \\
 & `INCONVENIENTES_AGLOMERACIONCASI NUNCA` + \\
 & TRATO_POR_CONDUCTORAGRESIVO + \\
 & TRATO_POR_CONDUCTORAMABLE + \\
 & TRATO_POR_CONDUCTORINDIFERENTE + \\
 & TIEMPO_ORIG_DESTINO10MIN + TIEMPO_ORIG_DESTINO15MIN + \\
 & TIEMPO_ORIG_DESTINO20MINOMAS + \\
 & TIEMPO_ORIG_DESTINO5MIN + `PORQUE_UTI_t.uBAJO COSTO` + \\
 & `PORQUE_UTI_t.uRAPIDEZ
 \end{aligned}$$

4.1.6 Validación del modelo

Según bibliografía es recomendable utilizar un modelo multinomial utilizando el intercepto, para poder comparar con los modelos anteriores y realizar la validación del modelo se utiliza la tabla cruzada de la variable endógena con las predicciones realizadas del modelo.

Tabla 4-13: Tasa de clasificación correcta e incorrecta del modelo 1

	BUENA	INSUFICIENTE	MALA
BUENA	123	14	7
INSUFICIENTE	35	69	45
MALA	14	30	58

Interpretación: La tabla 4-13 indica la tasa de clasificación correcta e incorrecta, donde la diagonal principal muestra la clasificación correcta, de ello se afirma que 14 personas que antes calificaron como buena ahora el modelo predijo como insuficiente, de la misma forma 35 personas que calificaron como insuficiente, el modelo indica que clasificaran como buena.

Clasificación correcta

```
sum(diag(res.cruz))/sum(res.cruz)
```

```
[1] 0.6329114
```

Clasificación incorrecta

```
1- sum(diag(res.cruz))/sum(res.cruz)
```

```
[1] 0.3670886
```

Pseudo R^2 de Mc-Fadden

```
rmcf <- 1-(deviance(res.mult1)/deviance(res.mult0))
rmcf
```

```
[1] 0.2874335
```

Pseudo R^2 de Cox-Shell

```
rcsh <- 1-exp((deviance(res.mult1)-deviance(res.mult0))/nrow(calida
```

```
d))
rcsh
[1] 0.4612968
```

Pseudo R^2 de Nagelkerke

```
rnk <- (1-exp((deviance(res.mult1)-deviance(res.mult0))/nrow(calida
d)))/(1-exp((-deviance(res.mult0))/nrow(calidad)))
rnk
[1] 0.5219694
```

El segundo modelo con las variables que son significativas en ambos casos

$$\begin{aligned}
 &VAL_TRANSPORTE \\
 &= TRATO_POR_CONDUCTORAGRESIVO \\
 &+ TRATO_POR_CONDUCTORAMABLE \\
 &+ TRATO_POR_CONDUCTORINDIFERENTE \\
 &+ TIEMPO_ORIG_DESTINO20MINOMAS \\
 &+ `PORQUE_UTI_t.uBAJO COSTO`
 \end{aligned}$$

Tabla 4-14: Tasa de clasificación correcta e incorrecta del modelo 2

	BUENA	INSUFICIENTE	MALA
BUENA	126	15	8
INSUFICIENTE	46	79	62
MALA	0	19	40

Interpretación: La tabla 4-14 indica la tasa de clasificación correcta e incorrecta, de modelo multinomial con las variables significativas, de ello se afirma que 15 personas que antes calificaron como buena ahora el modelo predijo como insuficiente, de la misma forma 46 personas que calificaron como insuficiente, el modelo indica que clasificaran como buena y 62 personas que antes calificaron como insuficiente el modelo predijo como mala.

Clasificación correcta

```
sum(diag(res.cruz1))/sum(res.cruz1)
[1] 0.6202532
```

Clasificación incorrecta

```
1- sum(diag(res.cruz1))/sum(res.cruz1)
[1] 0.3797468
```

Pseudo R^2 de Mc-Fadden

```
rmcf <- 1-(deviance(res.mult1)/deviance(res.mult0))
rmcf
[1] 0.2665322
```

Pseudo R^2 de Cox-Shell

```
rcsh <- 1-exp((deviance(res.mult1)-deviance(res.mult0))/nrow(calida
d))
rcsh
[1]0.4365117
```

Pseudo R^2 de Nagelkerke

```
rnk <- (1-exp((deviance(res.mult1)-deviance(res.mult0))/nrow(calida
d)))/(1-exp((-deviance(res.mult0))/nrow(calidad)))
rnk
[1] 0.4939243
```

Modelo logit ordinal

Aplico las mismas variables que predicen de mejor manera en el modelo multinomial, de acuerdo al modelo 2.

```
VVAL_TRANSPORTE
= TRATO_POR_CONDUCTORAGRESIVO
+ TRATO_POR_CONDUCTORAMABLE
+ TRATO_POR_CONDUCTORINDIFERENTE
+ TIEMPO_ORIG_DESTINO20MINOMAS
+ `PORQUE_UTI_t.uBAJO COSTO`
```

Tabla cruzada de clasificación

```
res.co <- table(predict(res.ord), calidad_dummy$VAL_TRANSPORTE)
```

Validación del modelo ordinal

Clasificación correcta

```
sum(diag(res.co))/sum(res.co)
```

```
[1] 0.6126582
```

Clasificación incorrecta

```
1- sum(diag(res.co))/sum(res.co)
```

```
[1] 0.3873418
```

Pseudo R^2 Mc-Fadden

```
ormf <- 1-(deviance(res.ord)/deviance(res.ord0))
```

```
ormf
```

```
[1] 0.2469706
```

Pseudo R^2 Cox-Snell

```
orch <- 1-exp((deviance(res.ord)-deviance(res.ord0))/nrow(calidad_d  
ummy))
```

```
orch
```

```
[1] 0.412283
```

Pseudo R^2 Nagelkerke

```
orn <- (1-exp((deviance(res.ord)-deviance(res.ord0))/nrow(calidad_d  
ummy)))/(1-exp((-deviance(res.ord0))/nrow(calidad_dummy)))
```

```
orn
```

```
[1] 0.466509
```

Resumen de resultados de la validación de modelos

Tabla 4-15: Tabla de resumen de los modelos elegidos

Modelo	Clasificación n correcta	Clasificación n Incorrecta	R^2	R^2	R^2
--------	-----------------------------	-------------------------------	-------	-------	-------

			de Mc-Fadden	de Cox-Snell	de Nagelkerke
1	0.6329	0.3671	0.2874	0.4613	0.5220
2	0.6203	0.3797	0.2665	0.4365	0.4939
3	0.6127	0.3873	0.2470	0.4123	0.4665

Interpretación: La tabla 4-15 se observa que el modelo 1 tiene 12 variables y un 63.29% de clasificación correcta y un 36.712% de incorrecta, seguidas por 28.74%, 46.13% y 52.20% de R_2 de MC.Fadden, Cox-Snell y Nagelkerke respectivamente lo cual indica que es un modelo y que explica de mejor manera a la variable Valoración de transporte.

El modelo 2 tiene 5 variables a pesar de eso consigue un total de clasificación correcta del 62.03% contra un total de 37.97% de clasificación incorrecta con un 26.65%, 43.65% y 49.39% de R_2 de MC.Fadden, Cox-Snell y Nagelkerke respectivamente lo cual indica que es un modelo y define bien la variable en estudio.

Por último el modelo 3 se basa en un logit ordinal con las 5 variables del modelo 2 y se obtiene una tasa de clasificación correcta del 61.27% contra un total de 38.73% de clasificación incorrecta con un 24.70%, 41.23% y 46.65% de R_2 de MC.Fadden, Cox-Snell y Nagelkerke respectivamente lo cual indica que es un modelo y define bien la variable en estudio.

Por el principio de parsimonia se elige el modelo logit multinomial (modelo 2), pues explica de mejor manera a la variable en estudio, además tiene una tasa de clasificación correcta mayor que el modelo ordinal, el modelo 1 es descartado pues contiene 12 variables con una diferencia de tasa de clasificación correcta del 1.27% con 7 variables más respecto al modelo 2.

Calculo de las ODD ratio del modelo logit Multinomial 2

Una vez determinado el modelo que más se ajuste a la evaluación de transporte, determinamos las ODD ratio, estas representan la probabilidad de ocurrencia de un evento frente a la no ocurrencia del evento.

Tabla 4-16: Cálculo de las ODD ratios del modelo final

CATEGORÍA	VARIABLE	OR	2.5 %	97.5 %
INSUFICIENTE	TRATO_POR_CONDUCTORAGRESIVO	1792,34	0	3,92E+20
	TRATO_POR_CONDUCTORAMABLE	0,01	0	6534,39
	TRATO_POR_CONDUCTORINDIFERENTE	0,17	0	100689,55
	TIEMPO_ORIG_DESTINO20MINOMAS	1,63	0,9	2,95
	PORQUE_UTI_t.uBAJO COSTO`	2,3	0,97	5,46
MALA	TRATO_POR_CONDUCTORAGRESIVO	3947,66	0	8,63E+20
	TRATO_POR_CONDUCTORAMABLE	0,01	0	3355,92
	TRATO_POR_CONDUCTORINDIFERENTE	0,13	0	76695,37
	TIEMPO_ORIG_DESTINO20MINOMAS	1,15	0,61	2,16
	PORQUE_UTI_t.uBAJO COSTO`	3,47	1,25	9,64

Interpretación: La evaluación de transporte se realiza en tres categorías: buena, insuficiente y mala, para la interpretación de las ODD ratios se tiene como p_1 , p_2 y p_3 respectivamente. Esto teniendo en cuenta que el modelo multinomial representa insuficiente/buena y mala/buena.

La evaluación de transporte urbano es insuficiente frente a una calificación buena para las personas que recibieron un trato por conductor agresivo en 1792.34 veces, indicando que si el conductor fue agresivo tiene mayor probabilidad de recibir una valoración insuficiente. Caso contrario cuando una persona reciba un trato amable la insatisfacción será de 0.01 veces lo que indica que mientras reciba un buen trato tienen menor probabilidad de recibir una mala calificación. Si recibe un trato indiferente el usuario puede valorar la calificación insuficiente en 0.17 veces. Un pasajero que pase en el servicio por más de 20 min tiene más probabilidad de dar una calificación insuficiente.

La evaluación del transporte urbano es mala frente a buena 3947.66 veces, esto indica que si el trato del conductor es agresivo es más probable que la valoración sea mala,

mientras que si el conductor brinda un trato amable es más probable que la valoración sea buena. Sin embargo, si el trato del conductor es indiferente la valoración es mala en 0.17 veces.

4.1.7 Presentación del modelo final

Como modelo más representativo se eligió el modelo logit multinomial 2, se plantea las ecuaciones a continuación:

Tabla 4-17: Coeficientes del modelo final

Notación	Variable	Insuficiente β_i	Mala β_j
	Intercepto	1,16	1,00
x_1	TRATO_POR_CONDUCTORAGRESIVO	7,49	8,28
x_2	TRATO_POR_CONDUCTORAMABLE	-4,54	-5,21
x_3	TRATO_POR_CONDUCTORINDIFERENTE	-1,79	-2,07
x_4	TIEMPO_ORIG_DESTINO20MINOMAS	0,49	0,14
x_5	PORQUE_UTI_t.uBAJO COSTO`	0,83	1,25

Interpretación: La tabla 4-17 contiene las variables significativas para el modelo y según los distintos Pseudo R^2 . Cada uno con su coeficiente. Para estudios futuros se debe considerar estas variables con más énfasis y de la misma forma las variables no repetitivas podrán ser descartadas y elaborar un nuevo cuestionario.

Para facilitar y simplificar el modelo denotamos como x_1, \dots, x_7

$$z_j = \frac{p_j}{p_{j \neq j}} = e^{\beta_{ij} + \beta_{i+1j}x_1 + \dots + \beta_{ij}x_p}$$

donde:

β_{ij} : son los coeficientes de la categoría j de la variable respuesta.

X_p : Variables exógenas que aportan al modelo.

$p_{j \neq j}$: Categoría frente a la cual se realiza el análisis.

Categorías de la variable endógena:

$p_{j=1}$: Buena

$p_{j=2}$: Insuficiente

$p_{j=3}$: Mala

Ecuación para la categoría insuficiente

$$z_2 = \frac{p_2}{p_1} = e^{\beta_{02} + \beta_{12}X_1 + \beta_{22}X_2 + \beta_{32}X_3 + \beta_{42}X_4 + \beta_{52}X_5}$$
$$z_2 = \frac{p_2}{p_1} = e^{1.16 + 7.49X_1 - 4.54X_2 - 1.79X_3 - 0.49X_4 + 0.83X_5}$$

Ecuación para la categoría mala

$$z_3 = \frac{p_3}{p_1} = e^{\beta_{03} + \beta_{13}X_1 + \beta_{23}X_2 + \beta_{33}X_3 + \beta_{43}X_4 + \beta_{53}X_5}$$
$$z_3 = \frac{p_3}{p_1} = e^{1 + 8.28X_1 - 5.21X_2 - 2.07X_3 + 0.14X_4 + 1.25X_5}$$

Para calcular las probabilidades de pertenecer a una categoría se realiza con las siguientes ecuaciones:

La probabilidad de que una persona califique el servicio como insuficiente es:

$$P(y = 2) = \frac{z_2}{1 + z_2 + z_3}$$

La probabilidad de que una persona califique el servicio como mala es:

$$P(y = 3) = \frac{z_3}{1 + z_2 + z_3}$$

Para calcular la probabilidad de que una persona califique el servicio como buena es el complemento de la probabilidad de ocurrencia de insuficiente y mala

$$P(y = 1) = 1 - [P(y = 2) + P(y = 3)]$$

Para entender el procedimiento realizamos un pronóstico para ello planteamos las ocurrencias de cada una de las variables dummy. Cabe recalcar que si el trato del conductor fue amable es remplazado por 1 mientras que el trato por conductor agresivo e indiferente serán 0, porque la misma persona no puede recibir los tres tratos del conductor.

Definimos como una persona que presento un trato agresivo, duro 20 min en el servicio y no viaja por bajo costo.

Procedemos con el cálculo de z_2 y z_3 , esto ayudara al cálculo de las probabilidades de valorar el servicio con respecto a una categoría: buena , insuficiente y mala.

$$z_2 = \frac{p_2}{p_1} = e^{1.16+7.49*1-4.54*0-1.79*0-0.49*1+0.83*0}$$

$$z_2 = \frac{p_2}{p_1} = 9351.84$$

$$z_3 = \frac{p_3}{p_1} = e^{1+8.28*1-5.21*0-2.07*0+0.14*1+1.25*0}$$

$$z_3 = \frac{p_3}{p_1} = 12381.82$$

Probabilidad de clasificar como insuficiente

$$P(y = 2) = \frac{9351.84}{1 + 9351.84 + 12381.82} = 0.4303 \approx 43.03\%$$

Existe un 43.03% de probabilidad de que el modelo clasifique como insuficiente cuando el pasajero viaje por más de 20 min y reciba un trato agresivo.

Probabilidad de clasificar como mala

$$P(y = 3) = \frac{12381.82}{1 + 9351.84 + 12381.82} = 0.5697 \approx 56.97\%$$

Existe un 56.97% de probabilidad de que el modelo clasifique como mala cuando el pasajero viaje por más de 20 min y reciba un trato agresivo.

Por tanto, la probabilidad de clasificar como buena es:

$$P(y = 1) = 1 - [0.4303 + 0.5697] = 0.0 \approx 0\%$$

Existe un 0% de probabilidad de que el modelo clasifique como buena el servicio de transporte, cuando el pasajero viaje por más de 20 min y reciba un trato agresivo.

Esta clasificación se corrobora con las ODD ratio, donde tenía más probabilidad de clasificar como mala o insuficiente mientras que reciba un trato agresivo

4.1.8 Contraste de razón de verosimilitudes del modelo multinomial

Planteamiento de hipótesis

H_0 : El modelo multinomial es significativo

H_1 : El modelo multinomial no es significativo

Nivel de significancia de $\alpha = 0,05$

Aplicación en Rstudio

```
anova(res.mult0, res.mult2)
```

donde:

res.mult0: representa al modelo nulo

re.mult2: representa al modelo multinomial 2 con 5 variables dummy

Tabla 4-18: Anova del contraste de razón de verosimilitudes

Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
1	788	850,08				

TRATO_POR_CONDUCTORAGRESIVO + TRATO_POR_CONDUCTORAMABLE + TRATO_POR_CONDUCTORINDIFERENTE + TIEMPO_ORIG_DESTINO20MINO MAS + PORQUE_UTI_tuBAJO COSTO	780	623,51	1 vs 2	8	226,574	0
--	-----	--------	--------	---	---------	---

Interpretación: La tabla 4-18 se valida el modelo significativo mediante una tabla ANOVA, para un nivel de significancia del 0.05 se contraste el test de razón de verosimilitudes para validar el modelo óptimo encontrado, se tiene un valor $p = 0$, lo que indica que el modelo 2 frente al modelo 0 es significativo y no rechaza la hipótesis nula, esto se concluye que las variables trato por conductor amable, agresivo, indiferente, tiempo de origen a destino de 20 minutos o más y la razón de uso del transporte urbano representan de manera óptima a la variable respuesta Valoración de transporte

Para un nivel de significancia del 0.05 se contraste el test de razón de verosimilitudes para validar el modelo óptimo encontrado, se tiene un valor $p = 0$, lo que indica que el modelo 2 frente al modelo 0 es significativo y no rechaza la hipótesis nula, esto se concluye que las variables trato por conductor amable, agresivo, indiferente, tiempo de origen a destino de 20 minutos o más y la razón de uso del transporte urbano representan de manera óptima a la variable respuesta Valoración de transporte

4.1.9 Análisis de la matriz pasajeros e Infraestructura de las vías (matriz 2)

La matriz contiene datos de los pasajeros que esperan en las distintas paradas en la ciudad, además si tiene señalización o no y el tipo de superficie por donde transitan cada uno de los transportes.

Datos faltantes o vacíos

El análisis de datos faltantes se las realiza por variable, esto con el fin de no descartar la información obtenida.

Tabla 4-19: Proporción de valores faltantes por variable

T_INICIO	T_FIN	LINEA
----------	-------	-------

0.00	0.00	0.00
PASAJEROS_SUBEN	PASAJEROS_BAJAN	PASAJEROS_NO_ACCEDEN
0.00	0.00	0.00
TIEMPO_ENTRE_PARADAS	ANCHO_CARRIL	N_CARRIL
0.00	0.17	0.17
TIPO_SUPERFICIE	TIPO_SEÑALIZACION_VIAL	ESTADO_SUP_RODADURA
1.75	0.00	0.17
PARADA	SEÑALIZACION_PARADA	
0.00	0.00	

Interpretación: la tabla 4-19 indica la proporción de datos faltantes donde las variables ancho de carril, numero de carril y estado de superficie de rodadura contiene un 0,17% de información faltante y el tipo de superficie de rodadura 1.71%, el tratamiento de estas se las realizara mediante la función *CentralImputation*.

Tabla 4-20: Imputación de datos faltantes

T_INICIO	T_FIN
0	0
LINEA	PASAJEROS_SUBEN
0	0
PASAJEROS_BAJAN	PASAJEROS_NO_ACCEDEN
0	0
TIEMPO_ENTRE_PARADAS	ANCHO_CARRIL
0	0
N_CARRIL	TIPO_SUPERFICIE
0	0
TIPO_SEÑALIZACION_VIAL	ESTADO_SUP_RODADURA
0	0
PARADA	SEÑALIZACION_PARADA

0	0
---	---

Interpretación: La tabla 4-20 indica el número de datos faltantes luego de realizar la imputación, donde se observó que no existen vacíos en las matrices, de la misma forma se imputa con la mediana y moda para las variable cuantitativas y cualitativas respectivamente.

4.1.10 Análisis descriptivo de Variables cualitativas

Tipo de superficie

Tabla 4-21: Frecuencias de la variable Tipo de superficie

Categoría	AD	ASF	HOR	LTR	PD
Frecuencia	15	414	135	3	6
Proporción	0.026	0.723	0.236	0.005	0.010

Interpretación: La tabla 4-21 contiene las frecuencias del tipo de superficie de ciudad, donde la mayor parte de estas es de asfalto seguido por hormigón, adoquín, piedra y lastre, cabe recalcar que cada uno de estos datos son tomados de los diferentes tramos por donde transitan los buses urbanos.

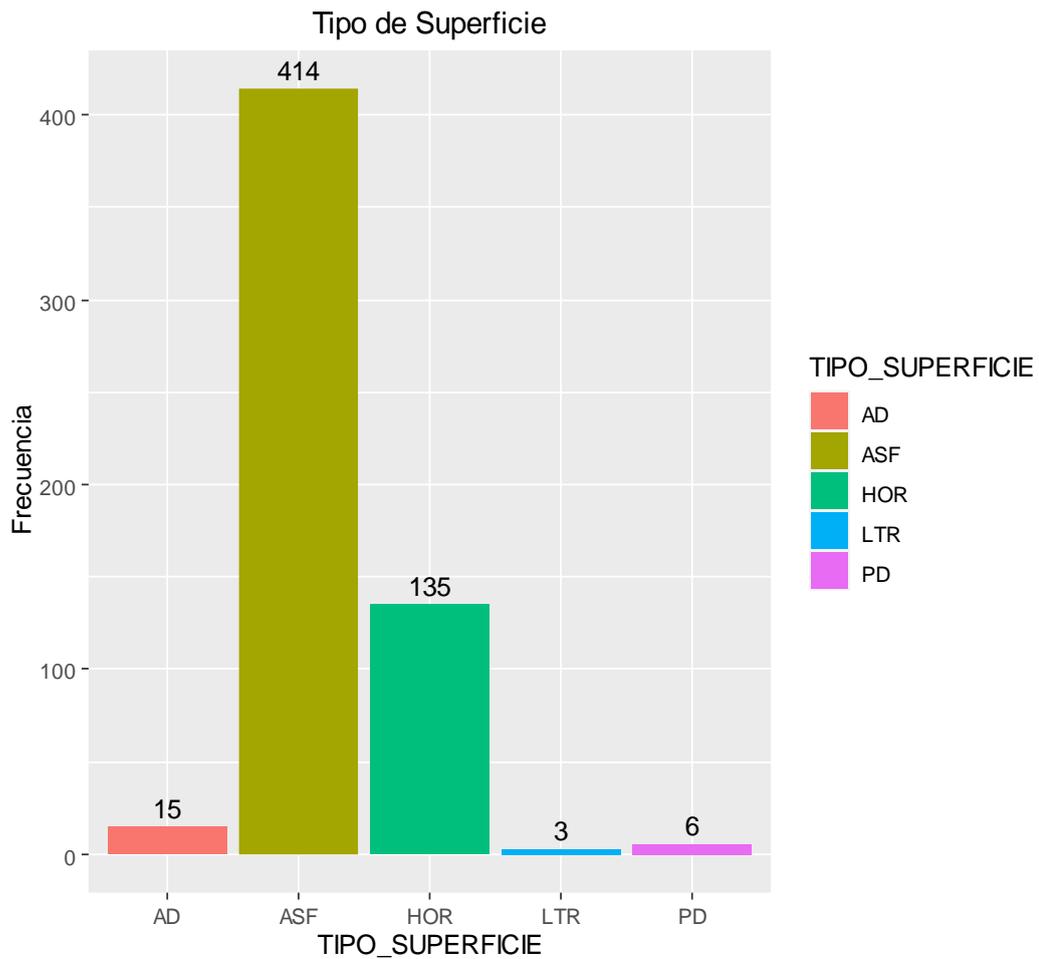


Figura 4-12: Frecuencias absolutas de la variable Tipo de superficie

El tipo de superficie donde transitan los buses urbanos es 15 tramos son superficie Adoquinado, 414 Asfalto, 135 de hormigón, 3 de lastre y 6 empedrado. Lo que es equivalente a 2.6%, 72,3%, 23.6%, 0.5% y 1% respectivamente.

Tipo de Señalización Vial

Tabla 4-22: Frecuencias de la variable Tipo de señalización vial

Categoría	H	H_V	SS	V
Frecuencia	69	290	109	105
Proporción	0.120	0.506	0.190	0.183

Interpretación: La tabla 4-22 contiene las frecuencias de los diferentes tipos de señal que cuentan cada uno de estos tramos, donde 69 cuentan con señalización horizontal, 290 con horizontal y vertical, 109 sin señalización, 105 con señalización vertical. Se debe tener en cuenta que la falta de señalización puede ocasionar accidentes.

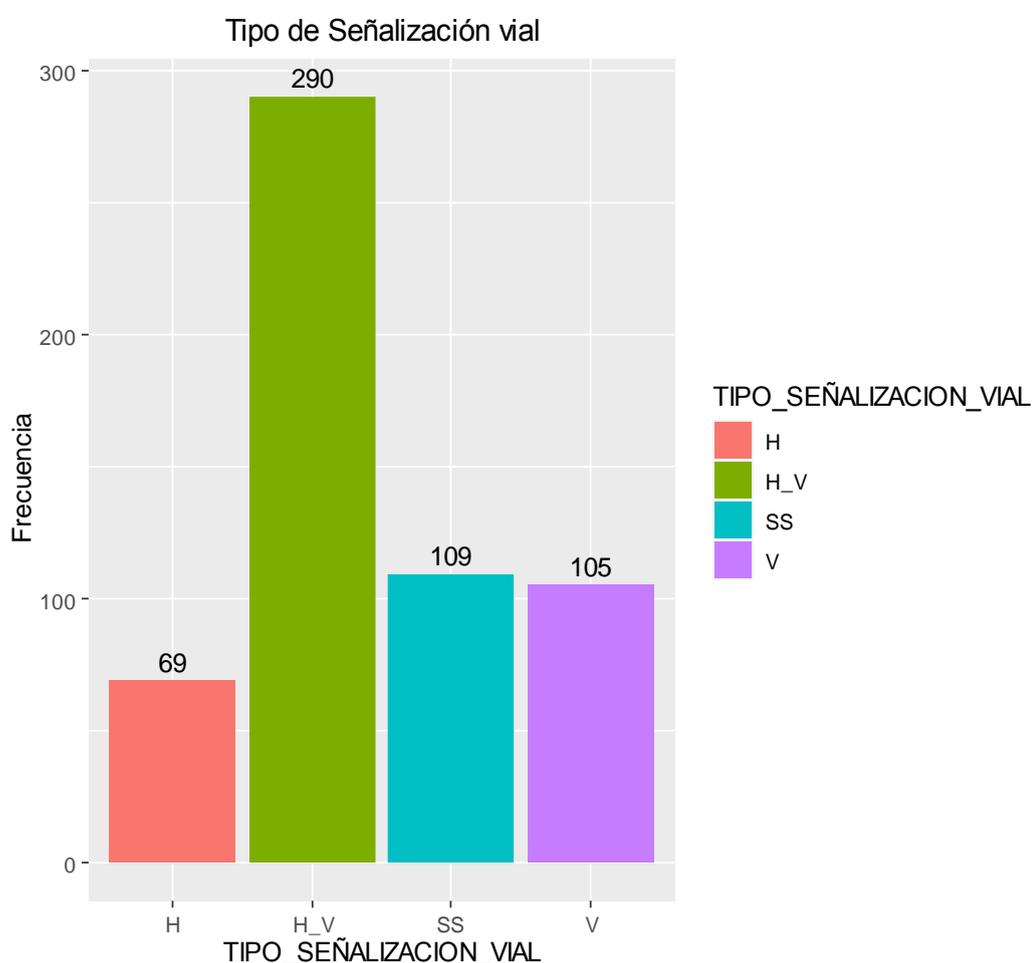


Figura 4-13: Frecuencias absolutas de la variable Tipo de señalización vial

Los diferentes tramos dentro de la ciudad cuentan 69 con señalización Horizontal, 290 con Horizontal y Vertical, 109 sin señalización y 105 con vertical. Esto representa

12%, 50.6%, 19% y 18,3% respectivamente.

Estado de Superficie de Rodadura

Tabla 4-23: Frecuencias de la variable Estado de superficie de rodadura

Categoría	B	E	M	R
Frecuencia	261	4	128	180
Proporción	0.455	0.007	0.223	0.314

Interpretación: La tabla 4-23 indica el estado de superficie de rodadura o por donde circula el transporte urbano, de las cuales 261 tramos están considerados en buen estado y solamente 4 en excelentes condiciones, pero 120 y 180 están consideradas en estado malo y regular respectivamente. Se debe realizar el respectivo mantenimiento pues una superficie de rodadura mala o regular afectara al rendimiento del bus urbano.

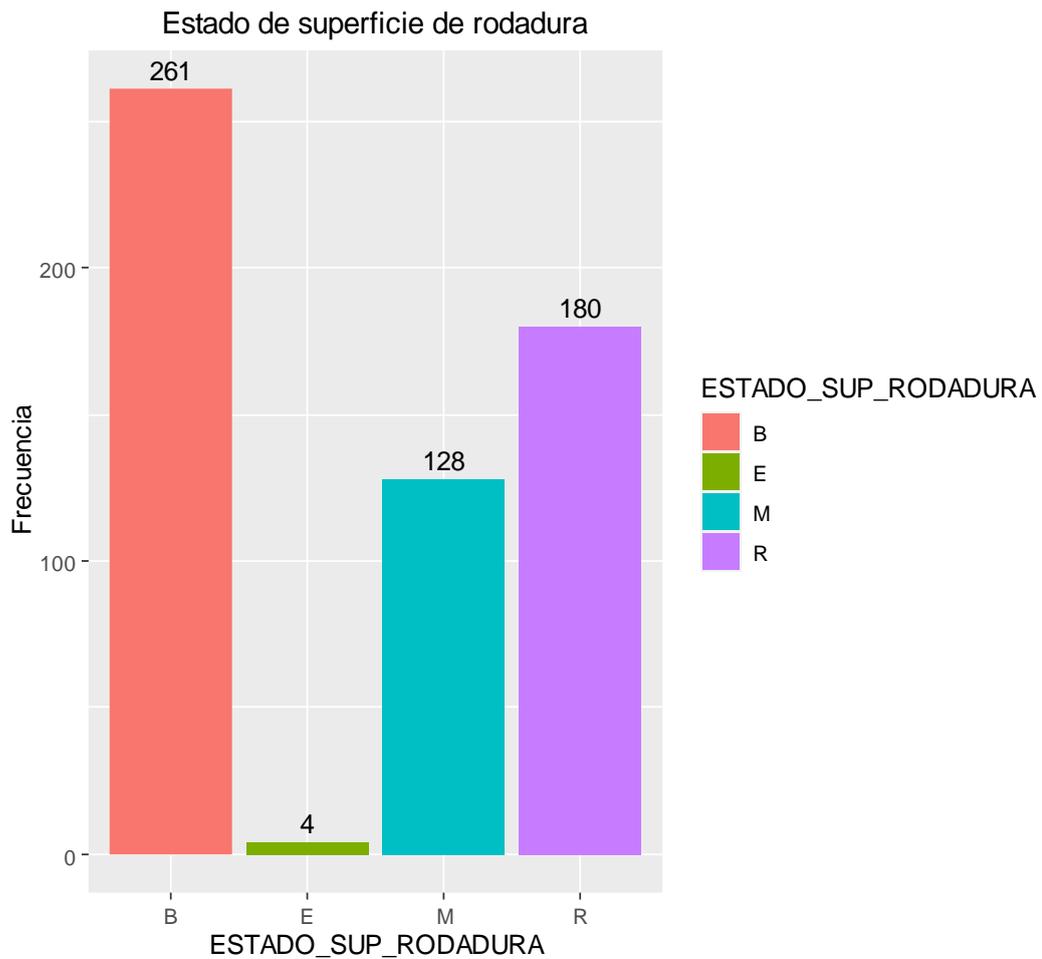


Figura 4-14: Frecuencias absolutas de la variable Estado de superficie de rodadura

En los diferentes tramos de la ciudad 261 presentan buen estado, 4 excelente, 128 mala y 180 Regular. Esto es equivalente al 45.5%, 0.7%, 22.3% y 31.4%.

Existencia de parada

Tabla 4-24: Frecuencias de la variable Existencia de parada

Categoría	E_P	N_EP
Frecuencia	176	397
Proporción	0.307	0.693

Interpretación: La tabla 4-24 muestra las frecuencias de la existencia de parada en cada uno de los tramos de la ciudad, donde existe una gran cantidad de paradas no establecidas o marcadas donde los usuarios puedan esperar el servicio. El gobierno o

la autoridad competente de la ciudad debe poner en marcha un proyecto de implementar paradas óptimas para los usuarios.

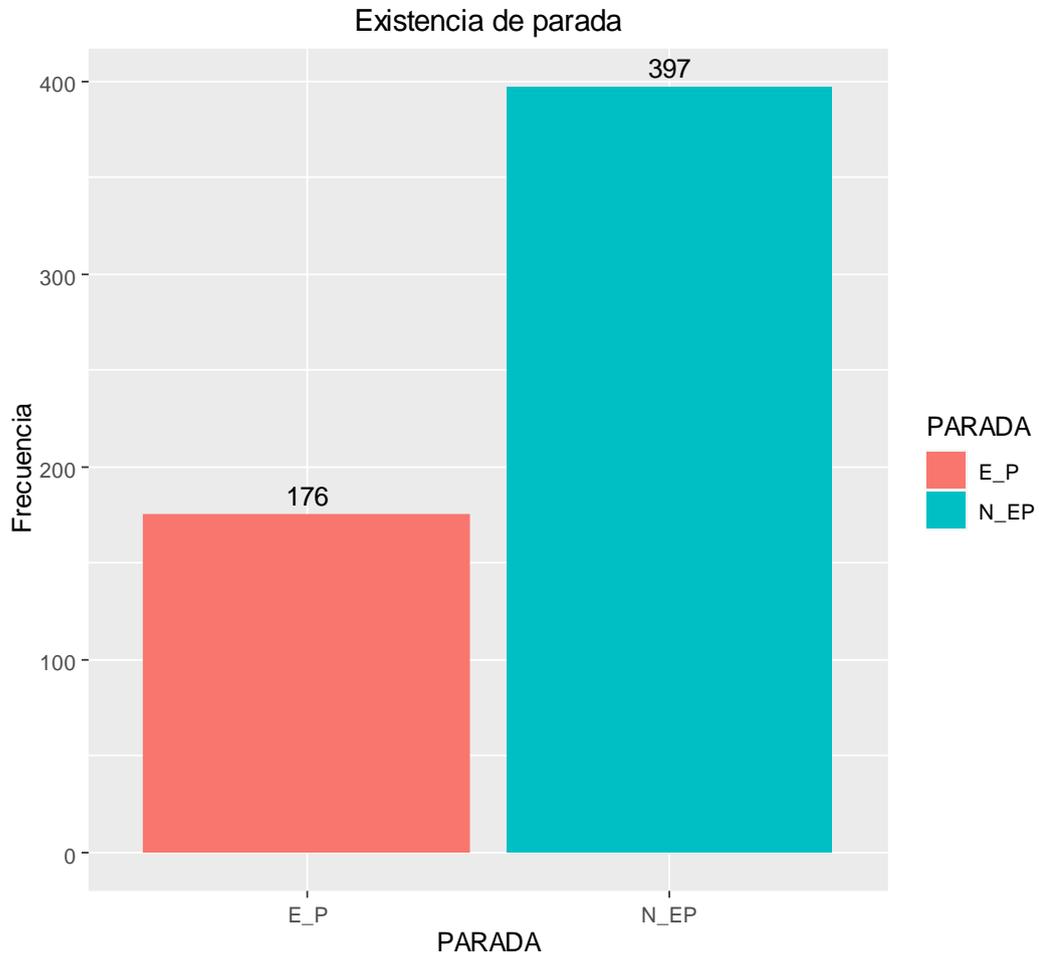


Figura 4-15: Frecuencias absolutas de la variable Existencia de parada

En la ciudad de Riobamba 176 tramos con parada y 397 tramos que no están con parada, lo que equivale a 30.7% y 69.3% respectivamente.

SEÑALIZACIÓN PARADA

Tabla 4-25: Frecuencias de la variable Señalización de parada

Categoría	H	P_H_V	S_S	V
Frecuencia	84	79	276	134
Proporción	0.147	0.138	0.482	0.234

Interpretación: La tabla 4-25 indica las frecuencias de señalización de paradas, pues en su gran mayoría las paradas no cuentan con una señalización y pocas cuentan con doble señalización, de estas la señalización vertical es más usada que las horizontales, pues se debería implementar los dos tipos para la comunicación entre conductores y peatones.

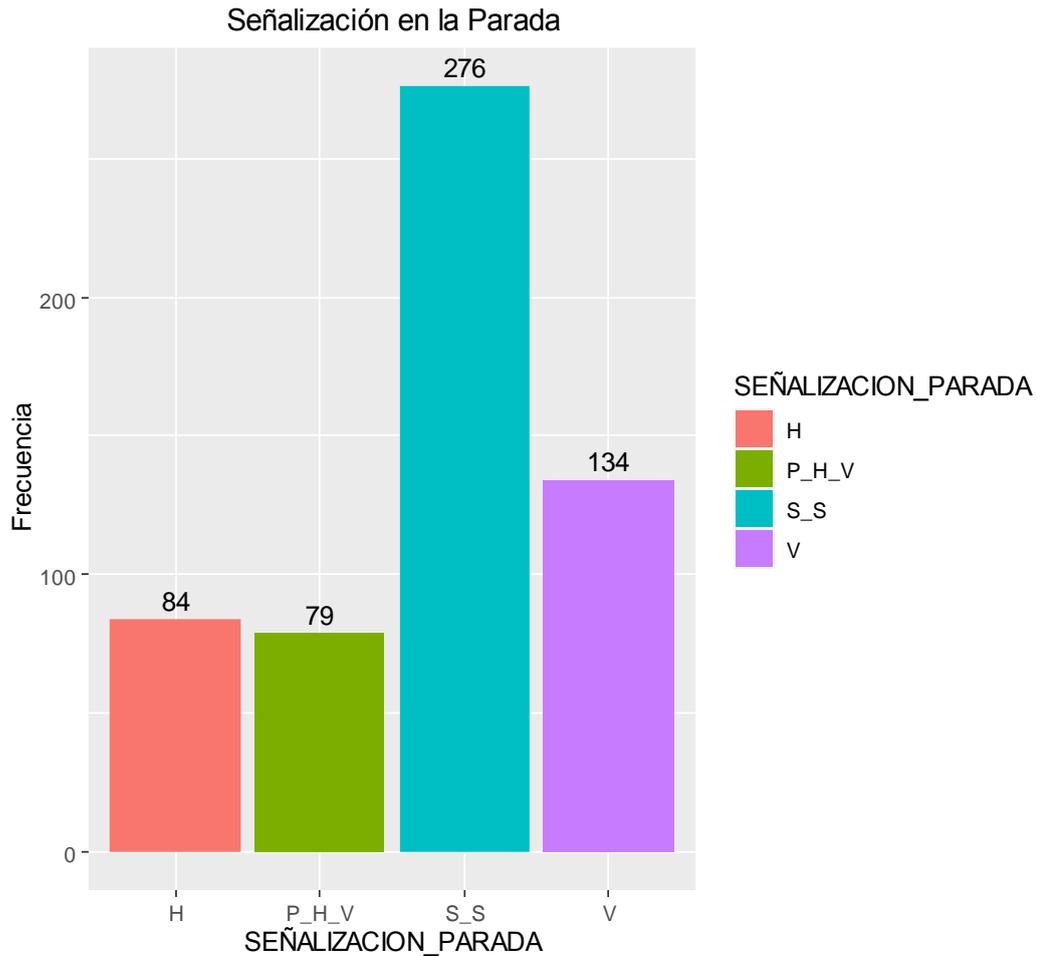


Figura 4-16: Frecuencias absolutas de la variable Señalización en parada

En los diferentes tramos de la ciudad 84 tienen señalización de parada horizontal, 79 Horizontal y vertical, 276 sin señalización y 134 con señalización vertical. Esto es equivalente a 14.7%, 13.8%, 48.2% y 23.4%.

4.1.11 Análisis descriptivo de variables cuantitativas

Media aritmética

Tabla 4-26: Media aritmética de variables cuantitativas

PASAJEROS_SUBEN	PASAJEROS_BAJAN
3.53	3.55
TIEMPO_ENTRE_PARADAS	ANCHO_CARRIL
2.73	3.90
PASAJEROS_NO_ACCEDEN	N_CARRIL
0.01	1.96

Interpretación: La tabla 4-26 indica la media aritmética de las variables cuantitativas, por tanto, el número de pasajeros promedio que suben o acceden al servicio de transporte es 3.53, el promedio de pasajeros que bajan en 3,55, el promedio de tiempo que demora es de 2,73 min, con un ancho de carril promedio de 3,90 metros, el promedio de pasajeros que no acceden es 0,01 y con un numero de carril 1,96 que se aproxima a 2 carriles.

Medidas de posición

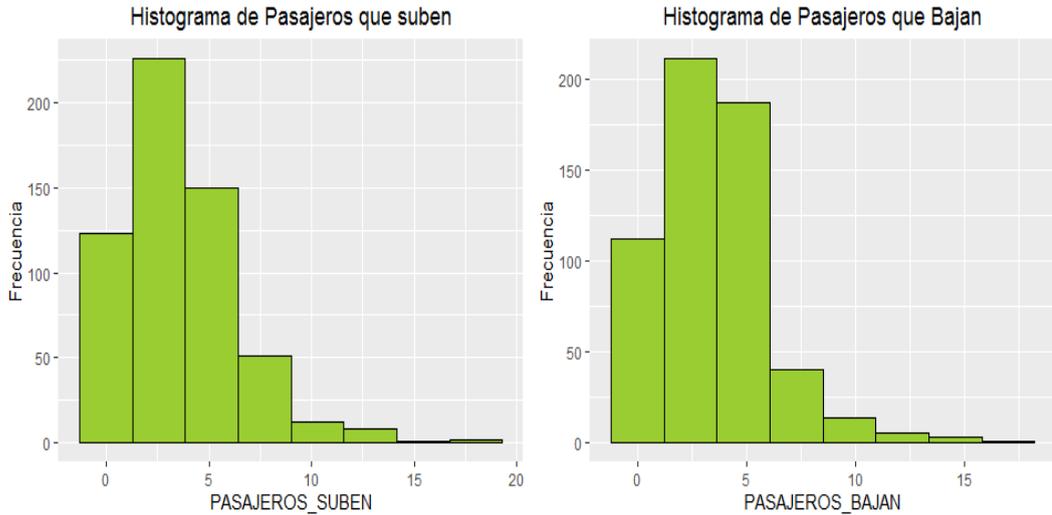
Tabla 4-27: Cuartiles por variable

Cuartil	25%	50%	75%	Interpretación
suben	2.0	3.0	5.00	El 25% de pasajeros que suben están menores o iguales a 2, el 50% es menor o igual a 3 y el 75% de la información esta menores o iguales a 5.
bajan	2.0	3.0	5.00	El 25% de pasajeros que bajan están menores o iguales a 2, el 50% es menor o igual a 3 y el 75% de la información esta menores o iguales a 5.
No acceden	0.0	0.0	0.00	El 25% de pasajeros que no acceden están menores o iguales a 0, el 50% es menor o igual a 0 y el 75% de la información esta menores o iguales a 0.
tiempo	2.0	3.0	3.00	El 25% de tiempo que se demora entre

				paradas están menores o iguales a 2, el 50% es menor o igual a 3 y el 75% de la información esta menores o iguales a 3.
Ancho carril	3.5	3.8	4.21	El 25% del ancho de carril están menores o iguales a 3.5, el 50% es menor o igual a 3.8 y el 75% de la información esta menores o iguales a 4.21.
# carril	2.0	2.0	2.00	El 25% de numero de carril están menores o iguales a 2, el 50% es menor o igual a 2 y el 75% de la información esta menores o iguales a 2.

Interpretación: La tabla 4-27 indican los valores correspondientes a los cuartiles, de ello se detalla el número de personas que están entre el 25%, 50 % y 75%.

Histogramas de variables cuantitativas



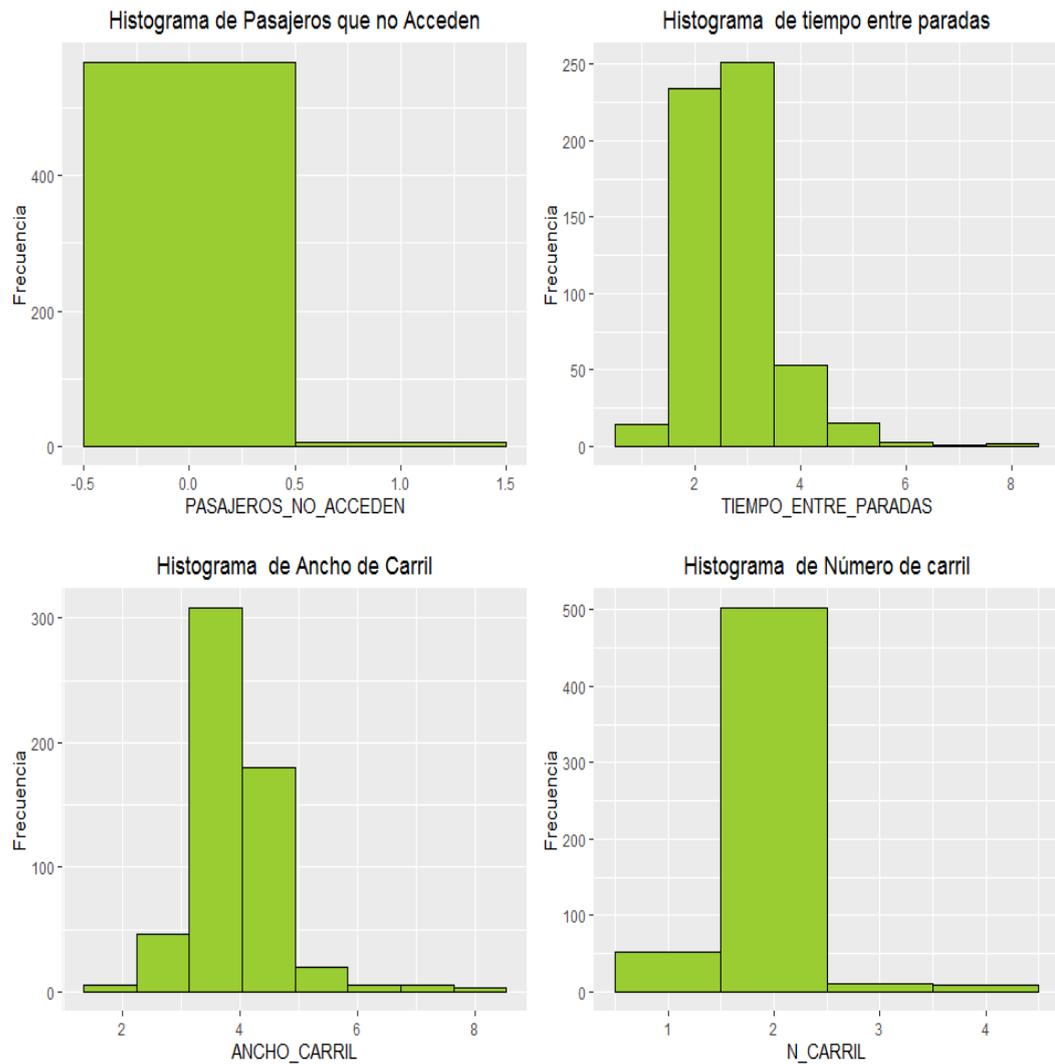


Figura 4-17: Histograma de variables cuantitativas

Observando los Histogramas notamos que el ancho de carril se aproxima a una curva normal mientras que las demás variables tienen asimetría positiva con cola a la derecha, la normalidad de las variables es un supuesto que se debe cumplir para el análisis de componentes principales y análisis factorial, La normalidad multivariada no se cumple tan solo una de las variables no cumple con esta condición.

4.1.12 Análisis del supuesto de Normalidad multivariada

Test de Shapiro Wilk de normalidad univariada

Tabla 4-28: Test de normalidad univariada

Test	Variable	Statistic	p value	Normality
------	----------	-----------	---------	-----------

Shapiro-Wilk	PASAJEROS_SUBEN	0.8640	<0.001	NO
Shapiro-Wilk	PASAJEROS_BAJAN	0.9025	<0.001	NO
Shapiro-Wilk	PASAJEROS_NO_ACEDEN	0.0747	<0.001	NO
Shapiro-Wilk	TIEMPO_ENTRE_PARADAS	0.7985	<0.001	NO
Shapiro-Wilk	ANCHO_CARRIL	0.8723	<0.001	NO
Shapiro-Wilk	N_CARRIL	0.4523	<0.001	NO

Interpretación: La tabla 4-28 contiene las pruebas de normalidad de las variables cuantitativas en estudio. Esto con el fin de determinar la normalidad multivariada supuesto importante para el análisis de componentes principales y Análisis factorial, donde para un nivel de significancia de 0.05 no se cumple la normalidad univariada, con ninguna de las variables, esto afecta a la normalidad multivariada.

Para un nivel de significancia de 0.05 no se cumple la normalidad univariada, con ninguna de las variables, esto será efecto en la normalidad multivariada.

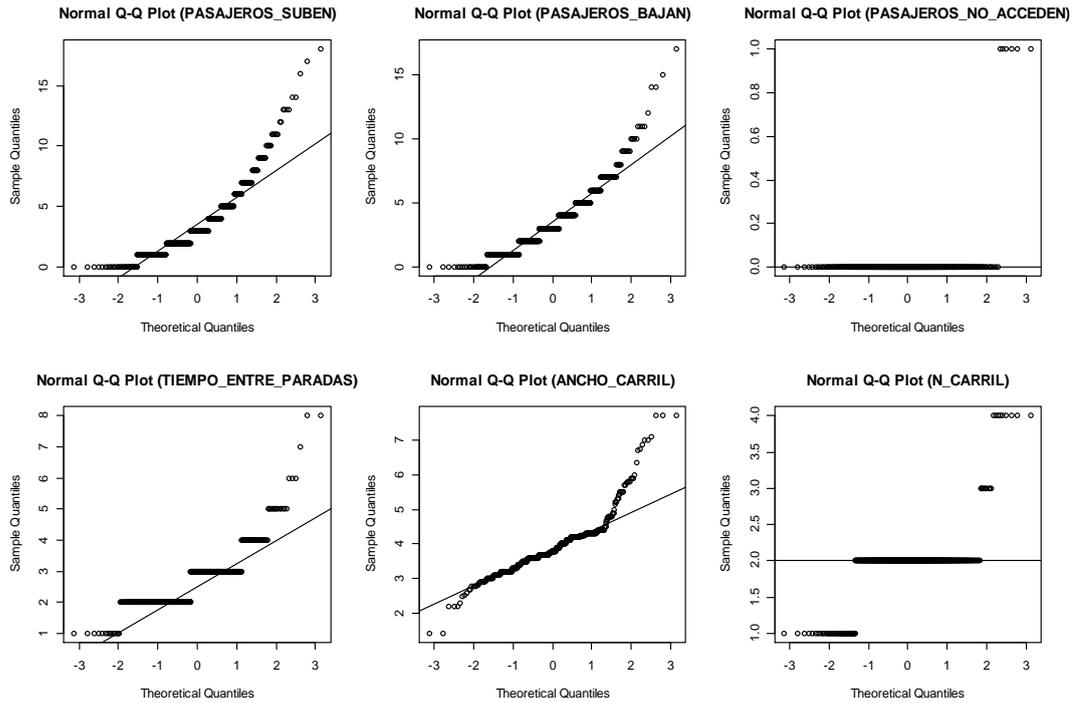


Figura 4-18: Q-Q plot de normalidad de las variables cuantitativas

La figura QQplot muestran la realidad de la información, donde es evidente que no cumplen la normalidad multivariada.

Test de Royston de normalidad Mutivariada

Tabla 4-29: Test de normalidad multivariada de Royston

Test	H	p value	MVN
Royston	743,9277	2E-157	NO

Interpretación: La tabla 4-29 realiza un test de hipótesis donde se prueba la normalidad multivariada y con un p valor menor al nivel de significancia de 0.05 se corrobora con lo analizado en la normalidad univariada, pues si una variable no cumpla con el supuesto afectara el cumplimiento en el análisis multivariado. Por tanto, rechazo la hipótesis nula y se concluye que la matriz de datos no es normal multivariado.

4.1.13 Análisis y componentes principales y análisis factorial

Para realizar el análisis de componentes principales y análisis factorial es primordial

que las variables tengan una alta correlación. Para lo cual analizamos la matriz de correlación y la prueba de esfericidad de Bartlett y la prueba de Kaiser-Meyer-Olkin. Esto se realiza como caso práctico pues las variables no cumplen la normalidad multivariada.

Matriz de correlación

Tabla 4-30: Matriz de correlaciones de las variables cuantitativas

	PASAJ EROS_ SUBEN	PASAJ EROS_ BAJAN	PASAJER OS_NO_A CCEDEN	TIEMPO_ ENTRE_P ARADAS	ANCH O_CA RRIL	N_C AR RIL
PASAJER OS_SUBE N	1	-0,25	0,12	0,33	0,05	- 0,03
PASAJER OS_BAJA N	-0,25	1	-0,03	0,06	-0,1	0
PASAJER OS_NO_A CCEDEN	0,12	-0,03	1	0,01	0,07	- 0,03
TIEMPO_ ENTRE_P ARADAS	0,33	0,06	0,01	1	0,14	- 0,08
ANCHO_C ARRIL	0,05	-0,1	0,07	0,14	1	- 0,25
N_CARRI L	-0,03	0	-0,03	-0,08	-0,25	1

Interpretación: La tabla 4-30 indica la correlación que existe entre las variables, donde se observó que la correlación más alta es de 0.33 entre pasajeros que suben y tiempo entre paradas mientras que la más baja es de 0 que corresponde a pasajeros que bajan, con una correlación negativa de -0.25 entre las variables pasajeros que suben y pasajeros que bajan.

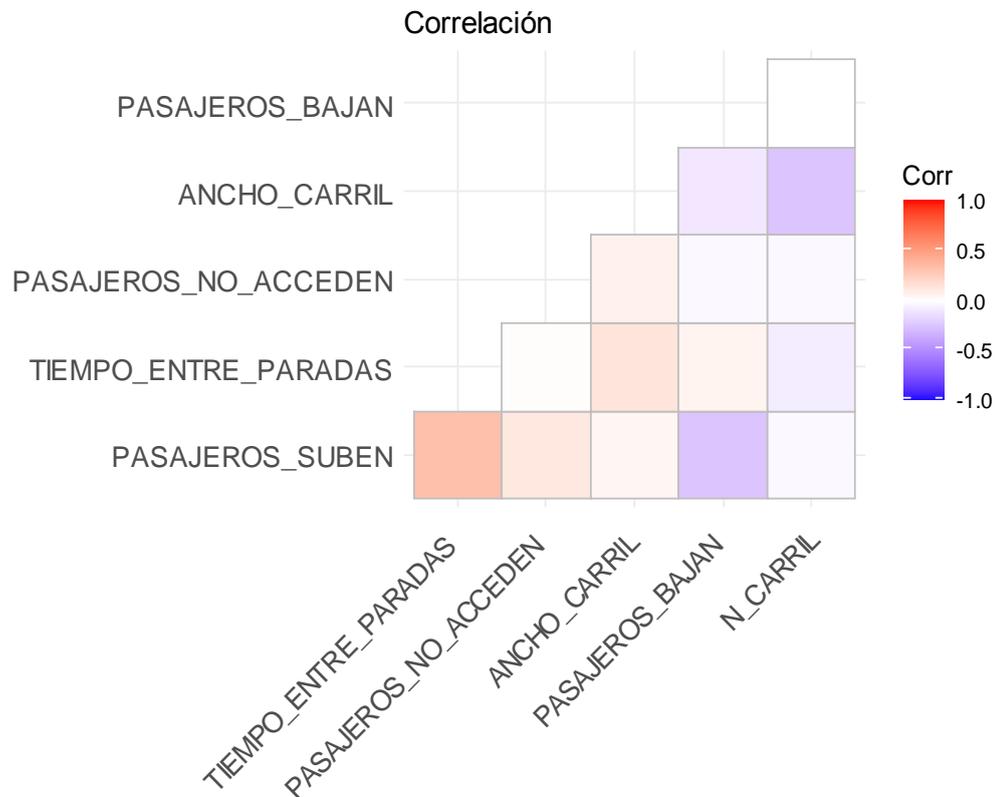


Figura 4-19: Figura de correlación de las variables numéricas

En la figura de correlación se observa por medio de colores el grado de asociación que tienen las variables, donde el azul claro indica una correlación negativa fuerte y el color rojo claro indica una correlación positiva fuerte.

Prueba de esfericidad de Bartlett

Verificar que la matriz sea factorizable.

Planteamiento de hipótesis

H₀: Las variables no están correlacionadas

H_a: Las variables están correlacionadas

```
p_esf <- cortest.bartlett(cor_usuario)
p_esf$p
```

```
[1] 0.00808497
```

Con un valor de $p = 0.008 < \alpha = 0.05$ no se rechaza H_0 por tanto se concluye que las variables en estudio no están correlacionadas.

La prueba de Kaiser-Meyer-Olkin

Los valores de KMO que estén se plantan de la siguiente manera

$KMO < 0.5$ no son aceptables

$KMO > 0.5$ son considerados aceptables

Si los valores del KMO son más próximos a 1 es más preciso el Análisis Factorial

```
KMO(cor_usuario)
```

```
Kaiser-Meyer-Olkin factor adequacy  
Call: KMO(r = cor_usuario)  
Overall MSA = 0.46
```

Se tiene un $KMO = 0.46$, esto indica que el análisis factorial, esto corrobora lo obtenido en la prueba de esfericidad de Bartlett y la matriz de correlaciones, pues se debe tener en cuenta que las variables deben estar correlacionadas de manera fuerte.

Scree plot

El scree plot es utilizada para analizar el número de factores o componentes principales óptimos que se deben incluir en un análisis, el grafico muestra los valores propios.

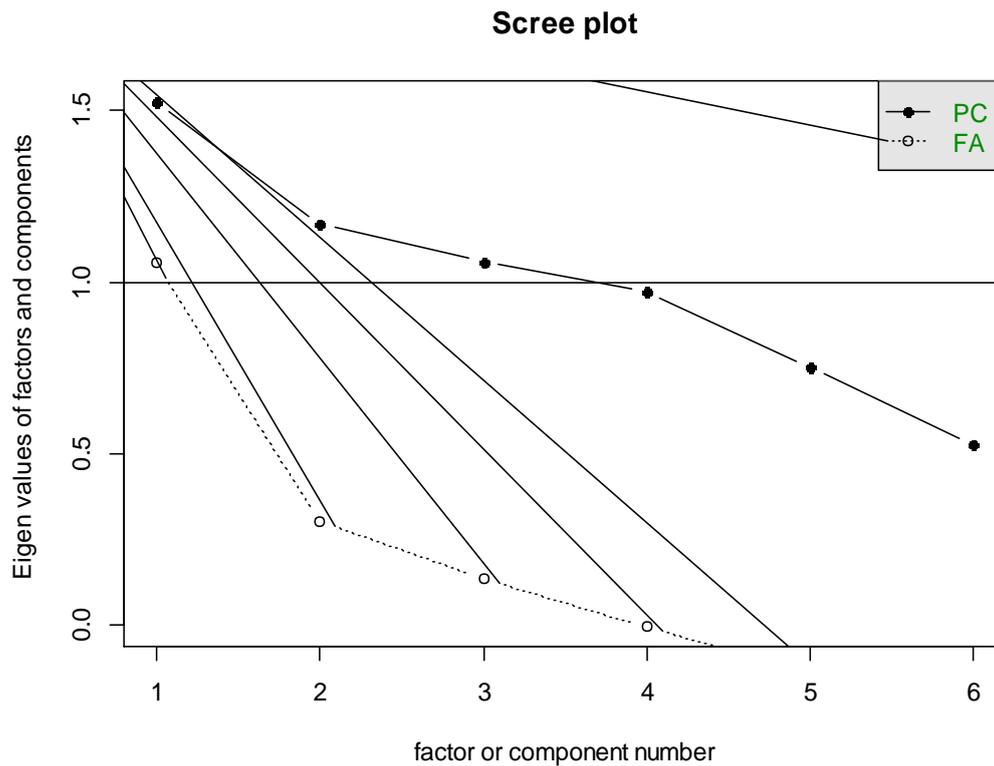


Figura 4-20: Scree plot de factores y componentes a utilizar

Según el scree plot solo se debe incluir un Factor lo que es inaceptable en el tipo de análisis esto es corroborado con el valor de KMO obtenido, mientras que para componentes principales es necesario utilizar las tres primeras componentes.

Componentes Principales

El análisis de componentes principales se realizará de acuerdo al scree plot analizado anteriormente donde proponía utilizar una cantidad de tres componentes.

Tabla 4-31: Coeficientes de cada componente

		PC1	PC2	PC3	PC4	PC5	PC6
PASAJEROS_SUBEN	Y_1	-0,57	0,45	0,14	-0,04	-0,22	-0,64
PASAJEROS_BAJAN	Y_2	0,30	-0,40	0,64	0,37	0,03	-0,45
PASAJEROS_NO_ACCEDEN	Y_3	-0,23	0,06	-0,27	0,92	-0,01	0,15

TIEMPO_ENTRE_PARADAS	Y_4	-0,49	0,04	0,65	-0,02	0,14	0,56
ANCHO_CARRIL	Y_5	-0,43	-0,48	-0,24	-0,10	0,67	-0,24
N_CARRIL	Y_6	0,32	0,63	0,11	0,09	0,69	-0,06

Interpretación: La tabla 4-31 indica los coeficientes de las combinaciones lineales de cada componente, donde la primera componente recoge mayor la información correspondiente a pasajeros que suben información seguido por tiempo entre paradas, de la misma forma la componente 2 recoge más la información de ancho de numero de carril seguido de ancho de carril, mientras que la componente 3 de pasajeros que bajan y pasajeros que no acceden.

$$PC1 = -0.57 * Y_1 - 0.40 * Y_4$$

$$PC1 = -0.48 * Y_5 + 0.63 * Y_6$$

$$PC1 = 0.64 * Y_2 - 0.27 * Y_2$$

Variabilidad explicada por cada componente

Tabla 4-32: Variabilidad explicada por cada componente

	PC1	PC2	PC3
Standard deviation	1,234975	1,079747	1,030334
Proportion of Variance	0,2542	0,1943	0,1769
Cumulative Proportion	0,25419	0,4485	0,62543

Interpretación: La tabla 4-32 indica la proporción explicada por cada componente, donde la primera componente principal explica el 25.42% de la variabilidad total, la segunda componente explica el 19.43% de la variabilidad total y la tercera componente explica el 17.69% de la variabilidad total. Las tres juntas explican el 62,54% la cual se considera aceptable de esto la dimensionalidad se reduce a 3. Teniendo en cuenta que las variables no cumplen con normalidad multivariada y el KMO no es aceptable, por

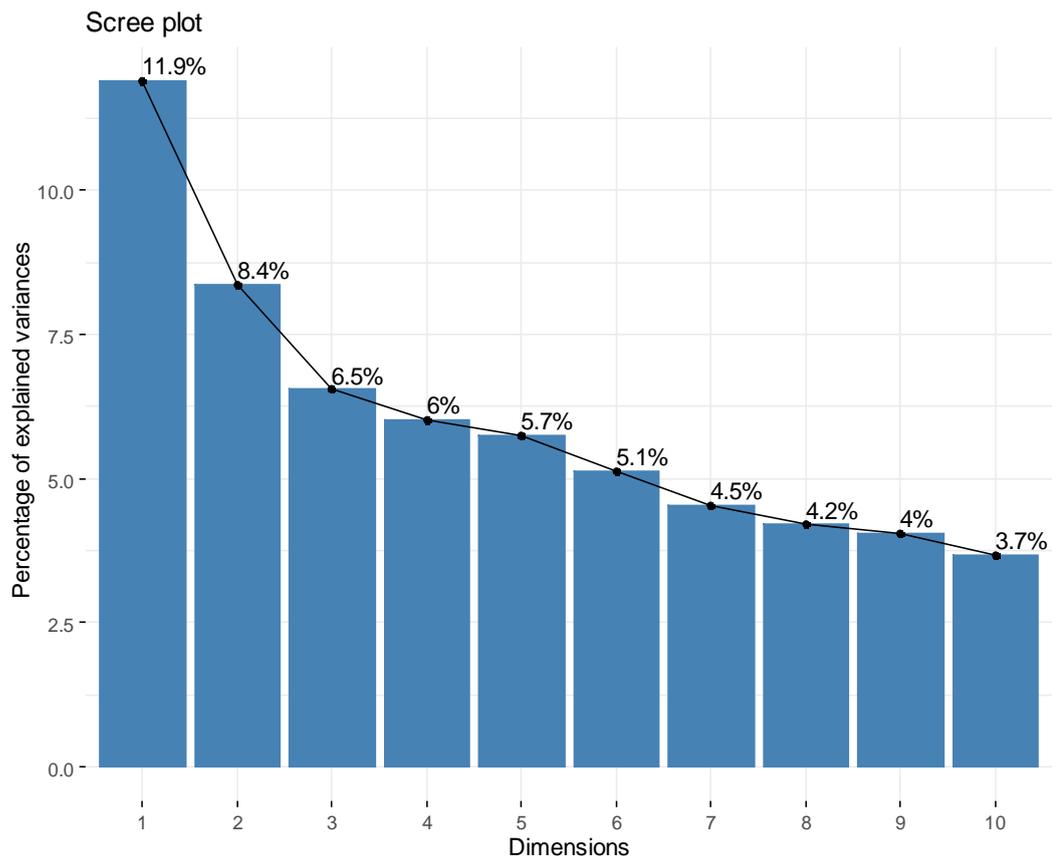


Figura 4-22: Porcentaje de variabilidad explicada por cada dimensión

El gráfico de sedimentación se la realizo para identificar el porcentaje de variabilidad explicada por cada dimensión, de la cual la primera representa el 11,9% de variabilidad total.

Correspondencia de las categorías

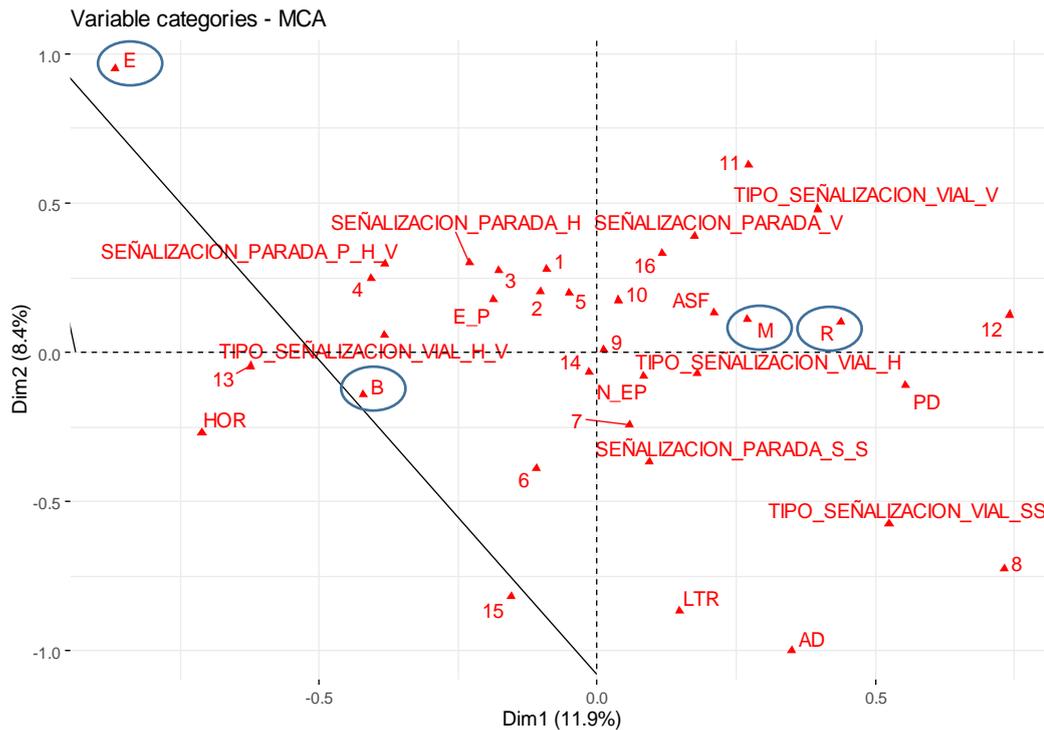


Figura 4-23: Correspondencia multivariada de las categorías

El hormigón una de las superficies que se considera en estado buena: la línea 4, 13 y 15 están más próximas a transitar por vías de buen estado además con señalizaciones de vías horizontales y verticales.

Las líneas 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 14 16, están más próximos de transitar por vías malas y regulares, así también existen tramos que no tienen parada ni señaléticas respectivas, de la misma forma no existe la señalización vial adecuada.

La línea 6, 8 y 15 transitan por tramos donde existen lastre y adoquinado, mientras que la línea 7 tiene más tramos sin señalización vial y parada.

Ninguna de las vías en los diferentes tramos está en estado excelente, sin embargo, si es de hormigón estará en excelente condición para el tránsito de los buses urbanos.

CAPÍTULO 5

5.1 Conclusiones

Las variables que mejor explican la valoración del transporte urbano fueron: el trato por conductor agresivo, amable e indiferente; el tiempo que dura un pasajero en el servicio de 20 minutos o más y la razón de uso del transporte urbano, ayudaron a una clasificación correcta del 62,03%. El modelo fue valorado mediante un R^2 de McFadden 0.2665, lo cual es considerado como un modelo aceptable esto se corrobora con el R^2 de Cox-Shell y Nagelkerke de 0.4365 y 0.4939 respectivamente.

El determinar estudios breves respecto al transporte público puede caer en usar métodos poco apropiados. Para realizar un “Análisis de Componentes Principales” se recomienda realizar previamente una prueba de Kaiser-Meyer-Olkin (KMO), dado que cualquier software le va a permitir realizar el ACP pero me puede dar resultar nefastos si no existe una relación alta entre las variables

Otro aspecto importante para la valoración del transporte público de pasajeros fue utilizar el análisis de correspondencia, el cual nos indicó que muchos de las líneas de transporte circulan por vías en estado regular y malo, de la misma forma las señaléticas viales y de parada no existen en algunos tramos.

La información que se obtuvo para la investigación generado por tres aspectos importantes dentro del transporte público de pasajeros como son infraestructura, ascenso y descenso de pasajeros y calidad del servicio está estructurada en una base de treinta y seis mil datos que puedes servir para otro tipo de investigaciones e incluso otros modelos.

5.2 Recomendaciones

La valoración de transporte urbano en la ciudad de Riobamba es muy importante pues de esta forma las diferentes entidades que brindan este servicio podrán mejorar su calidad de servicio, además el estado de las vías de circulación es importante puesto que el bus urbano se mantendrá en mejores condiciones que se adecuen a la perspectiva del usuario. Dicho esto, es recomendable analizar nuevas formas de medir la calidad del servicio en esta ciudad.

Se recomienda para el levantamiento de información respecto al transporte, se realice un pilotaje al cuestionario a ser utilizado, o a su vez se siga matrices ya probadas en este campo.

Finalmente, se recomienda que al realizar estudios o investigaciones referente a la valoración del transporte o medir la calidad del servicio del transporte, no solo se realice un análisis descriptivo del mismo, sino que se realice un modelado con la información obtenida para tener mejores apreciaciones y tomar acertadas decisiones.

5.3 BIBLIOGRAFÍA

- 1 Brahmaiah, B., Prasad, A., & Srinivas, K. (2017). A Performance Analysis Of Modelling Route Choice Behavior On Urban Bus. *International Journal of Advanced Information Science and Technology*, Vol. 6(N° 1), 1-10. doi:10.15693/ijaist/2017.v6i1.1-10
- 2 Chowdhury, S., & Van Wee, B. (2020). Examining women's perception of safety during waiting times at public transport terminals. *Transport Policy*, 94, 102-108. doi:<https://doi.org/10.1016/j.tranpol.2020.05.009>
- 3 Delgado, H., Prado, V., & Ochoa, J. (s.f). ESTUDIO DE LA CALIDAD DEL SERVICIO DE TRANSPORTE URBANO EN CD. GUZMAN, JALISCO (DESDE LA PERCEPCION DEL USUARIO). (U. d. Guadalajara, Ed.) *Investigación de Ciencias Administrativas*. Obtenido de http://acacia.org.mx/busqueda/pdf/01_04_Calidad_del_Servicio.pdf
- 4 Delgado, M., Gómez, A., & De Esteban, J. (2019). The social perception of urban transport in the city of Madrid: the application of the Servicescape Model to the bus and underground services. *Springer Link*(37). doi:<https://doi.org/10.1186/s12544-019-0373-5>
- 5 Duarte, A., Mendoza, Y., Arévalo, A., & Guerrero, T. (2015). MODELACIÓN DE LA ELECCIÓN MODAL EN TRANSPORTE PÚBLICO URBANO A PARTIR DE LA INTRODUCCIÓN DE UN NUEVO MODO. *Revista Colombia de Tecnologías de Avanzada*, Vol. 1(N° 25). doi:<https://doi.org/10.24054/16927257.v25.n25.2015.2369>
- 6 Fajardo, C., & Gómez, A. (2015). Análisis de la elección modal de transporte público y privado en la ciudad de Popayán. *Territorios*.
- 7 Garcia, P. (2005). Una aproximación microeconómica a los determinantes de la elección del modo de transporte. *Revista Latinoamericana de Desarrollo Económico*. doi:2074-4706
- 8 José, L., & Gustavo, A. (2019). Evaluación de la calidad del servicio de transporte urbano en bus de la ciudad de Ambato. *Visionario Digital*, Vol. 3 (N° 2), 26-46. Recuperado el 17 de 11 de 2020, de <https://doi.org/10.33262/visionariodigital.v3i2.392>
- 9 Llanos, Y., Llanos, D., & Molina, I. (2013). PERCEPCIÓN DE LA CALIDAD DEL SERVICIO DE TRANSPORTE URBANO EN LA CIUDAD DE FLORENCIA. *evista Estrategia Organizacional*. Recuperado el 17 de 11 de 2020, de https://www.researchgate.net/publication/319149778_PERCEPCION_DE_LA_CALIDAD_DEL_SERVICIO_DE_TRANSPORTE_URBANO_EN_LA_CIUDAD_DE_FLORENCIA/fulltext/5994fc7aa6fdccaded2109e1/PERCEPCION-DE-LA-CALIDAD-DEL-SERVICIO-DE-TRANSPORTE-URBANO-EN-LA-CIUDAD-DE-FLORENCIA.
- 10 Orellana, D. (2016). Métodos para el análisis de patrones de movilidad no motorizada. *Universidad de Cuenca*. Recuperado el 16 de 11 de 2020, de <https://lactalab.ucuenca.edu.ec/wp-content/uploads/2018/02/Orellana-Metodos-movilidad-no-motorizada-2016.pdf>
- 11 Padilla, M., Llamuca, J., Calderón, P., & Villamarin, J. (2019). Modelos matemáticos y la calidad del servicio aplicados al transporte urbano en Riobamba. *CIENCIA DIGITAL*, Vol. 3(2.2), 146-160. Recuperado el 17 de 11 de 2020, de <https://cienciadigital.org/revistacienciadigital2/index.php/CienciaDigital/article/view/466>
- 12 Pando, V., & San Martin, R. (2004). Regresion Logistica Multinomial. doi:1575-2410
- 13 Roche, I., Rundmo, T., Foss, J., & Moe, D. (2013). Transport mode preferences, risk perception and worry in a Norwegian urban population. *Accident Analysis & Prevention*, 50, 698-704. doi:<https://doi.org/10.1016/j.aap.2012.06.020>
- 14 Rodríguez, L., Chacón, E., & Orozco, E. (2018). Percepción de la calidad del servicio en el sistema de transporte unificado (STU) de Santa Marta, Colombia. *Revista Espacios*, Vol. 39(N° 49), 35-49. Recuperado el 16 de 11 de 2020, de <http://www.revistaespacios.com/a18v39n47/a18v39n47p35.pdf>

- 15 Romero, G., & Lugo, D. (2018). El estado del arte de la movilidad del transporte en la vida urbana en ciudades latinoamericanas. *Revista transporte y territorio*, 133-157. doi:1852-7175
- 16 Romero, J., Sánchez, N., & Iara, E. (2016). Percepción de la satisfacción del servicio en el transporte público solo para mujeres. *Revista Transporte y Territorio*, 164-182. doi:1852-7175
- 17 Sánchez, O., & Romero, j. (2010). Factores de calidad del servicio en el transporte público de pasajeros: estudio de caso de la ciudad de Toluca, México. *SciCIELO*. doi:1405-8421
- 18 Vega, O., Rivera, H., & Malaver, N. (2017). Contrastación entre expectativas y percepción de la calidad de servicio del sistema de transporte público de autobuses en Bogotá. *Revista Espacios*, Vol. 38(Nº 43). Recuperado el 16 de 11 de 2020, de <https://www.revistaespacios.com/a17v38n43/17384303.html>

5.4 ANEXOS

Base infraestructura

	SERVIDOR	LINEA	INICIAL	FINAL	CANTIDAD	PLAZA	TIPO	INICIAL	FINAL	INICIAL	FINAL	INICIAL	FINAL	INICIAL	FINAL	TIEMPO		PASAJEROS QUE SUBEN	PASAJEROS QUE BAJAN	PASAJEROS QUE NO	TIEMPO ENTRE	
																INICIO	FIN					
1																						
2																						
3																						
4																						
5																						
6																						
7																						
8																						
9																						
10																						
11																						
12																						
13																						
14																						
15																						
16																						
17																						
18																						

Base Calidad

	PRESTACION	FECHA	USUARIO	SENERO	NOTA_TRANSPORTE	TIEMPO ESPERA	DEFINIDA PASAJEROS	INCÓMODO/AS	ASIGNACIONES	CAUSACIONES	NOTAS	TIEMPO OMO	ESTADO	TRATOS	POSA	CONDICIONES	DE SEGURIDAD	PORQUE	USU	LA	
1																					
2																					
3																					
4																					
5																					
6																					
7																					
8																					
9																					
10																					
11																					
12																					

Base Asenso – Descenso

	SERVIDOR	LINEA	INICIAL	FINAL	CANTIDAD	PLAZA	TIPO	INICIAL	FINAL	INICIAL	FINAL	INICIAL	FINAL	INICIAL	FINAL	TIEMPO		PASAJEROS QUE SUBEN	PASAJEROS QUE BAJAN	PASAJEROS QUE NO	TIEMPO ENTRE	
																INICIO	FIN					
1																						
2																						
3																						
4																						
5																						
6																						
7																						
8																						
9																						
10																						
11																						
12																						
13																						
14																						
15																						
16																						
17																						
18																						
19																						
20																						
21																						
22																						
23																						
24																						
25																						
26																						
27																						
28																						
29																						
30																						

Datos

NUMERO BU	KM INICIAL	KM FINAL	HORA SALIDA	HORA LLEGA	LINEA	TRAMO		PASAJEROS QUE SUBEN	PASAJEROS QUE BAJAN	PASAJEROS QUE NO	TIEMPO ENTRE
						INICIO	FIN				
130	258697	258881	6:13	7:36	1	1	2	15	1	0	6:15
130	258697	258881	6:13	7:36	1	2	3	1	1	0	6:16
130	258697	258881	6:13	7:36	1	3	4	35	0	0	6:20
130	258697	258881	6:13	7:36	1	4	5	1	0	0	6:21
130	258697	258881	6:13	7:36	1	5	6	3	0	0	6:23
130	258697	258881	6:13	7:36	1	6	7	0	3	0	6:24
130	258697	258881	6:13	7:36	1	7	8	1	0	0	6:24
130	258697	258881	6:13	7:36	1	8	9	1	2	0	6:25
130	258697	258881	6:13	7:36	1	9	10	4	2	0	6:28
130	258697	258881	6:13	7:36	1	10	11	0	2	0	6:30
130	258697	258881	6:13	7:36	1	11	12	0	5	0	6:32
130	258697	258881	6:13	7:36	1	12	13	11	8	0	6:35
130	258697	258881	6:13	7:36	1	13	14	4	4	0	6:37
130	258697	258881	6:13	7:36	1	14	15	0	2	0	6:39
130	258697	258881	6:13	7:36	1	15	16	0	1	0	6:41
130	258697	258881	6:13	7:36	1	16	17	1	6	0	6:44
130	258697	258881	6:13	7:36	1	17	18	2	22	0	6:46
130	258697	258881	6:13	7:36	1	18	19	3	3	0	6:48
130	258697	258881	6:13	7:36	1	19	20	0	3	0	6:50
130	258697	258881	6:13	7:36	1	20	21	0	1	0	6:51
130	258697	258881	6:13	7:36	1	21	22	7	3	0	6:55
130	258697	258881	6:13	7:36	1	22	23	15	9	0	6:58
130	258697	258881	6:13	7:36	1	23	24	9	2	0	7:00
130	258697	258881	6:13	7:36	1	24	25	4	12	0	7:02
130	258697	258881	6:13	7:36	1	25	26	20	4	0	7:06