# Bivariate analysis of the Bogota metro project

Author: Chuwen Zhong, Zhanchao Yang </br>

Summer 2025

## Purpose and Method of Bivariate Analysis

### Why We Are Doing This Bivariate Analysis

The purpose of the bivariate analysis in this study is to:

- **Explore associations** between pairs of variables from the Bogotá Metro survey, such as:

- Public perception of the Metro vs housing characteristics

- Housing tenure vs commuting behavior

- Income vs willingness to pay for Metro fare

- Current transportation mode vs willingness to adopt Metro

- **Test hypotheses** such as:

- Renters are more optimistic about the Metro than owners.

- Renters prefer living near mass transit to reduce commute time.

- Higher-income households are more willing to pay higher fares.

- Current transit users are more likely to use the Metro.

- **Inform regression modeling** by:

- Revealing trends and potential predictor-outcome relationships

- Identifying sparse or small cell sizes that may require category collapsing

- Assessing whether data meet assumptions for regression models

### Method

The analysis uses **cross-tabulation (contingency tables)** to show how the values of one variable are distributed across another. Specifically, the methods include:

- **Frequency and percentage tables** for each relationship

- **Stratification by Metro Line** (Line 1 and Line 2) to compare geographic differences

- **Descriptive summaries** of associations and patterns in the tables

These methods are **non-parametric and exploratory**, serving as a foundation for the proposed regression models, such as:

- Ordinal logistic regression (e.g., perception or willingness as outcomes)

- Binary logistic regression (e.g., willingness to use the Metro: yes/no)

- Multinomial logistic regression (if the proportional-odds assumption is violated)

## Relationship 1: public perception vs housing characteristics

### Research Question

**Perception regarding the construction and functioning of the first and Second metro Lines of Bogota vs residents' housing characteristics**

The survey question involved `P67` vs `P82, P83, P86`

### Hypothesis:

- The more optimistic the person is, the higher the probability that they are a renter

- The more pessimistic the person is, the higher the probability that the person is an owner

### Bivariate table

#### Perception vs household ownership

**Line 1: Perception vs household ownership**

|  |  |  |  |
| --- | --- | --- | --- |
| P67 | Own | Rent | Total |
| 1 (Very pessimistic) | 4.7% (7) | 7.1% (10) | 5.9% (17) |
| 2 | 6.1% (9) | 5.0% (7) | 5.5% (16) |
| 3 | 21.6% (32) | 14.9% (21) | 18.3% (53) |
| 4 | 17.6% (26) | 17.7% (25) | 17.6% (51) |
| 5 | 16.9% (25) | 18.4% (26) | 17.6% (51) |
| 6 (Very optimistic) | 33.1% (49) | 36.9% (52) | 34.9% (101) |
| \*\*Total\*\* | \*\*100.0% (148)\*\* | \*\*100.0% (141)\*\* | \*\*100.0% (289)\*\* |

**Line 2: Perception vs household ownership**

|  |  |  |  |
| --- | --- | --- | --- |
| P67 | Own | Rent | Total |
| 1 (Very pessimistic) | 16.2% (12) | 12.7% (22) | 13.8% (34) |
| 2 | 8.1% (6) | 3.5% (6) | 4.9% (12) |
| 3 | 14.9% (11) | 24.9% (43) | 21.9% (54) |
| 4 | 18.9% (14) | 19.7% (34) | 19.4% (48) |
| 5 | 16.2% (12) | 16.8% (29) | 16.6% (41) |
| 6 (Very optimistic) | 25.7% (19) | 22.5% (39) | 23.5% (58) |
| \*\*Total\*\* | \*\*100.0% (74)\*\* | \*\*100.0% (173)\*\* | \*\*100.0% (247)\*\* |

#### Perception vs living time

**Line 1: Perception vs living time**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P67 | ≤ 1 year | > 1 and ≤ 5 years | > 5 and ≤ 10 years | > 10 and ≤ 15 years | > 15 and ≤ 20 years | > 20 years | Total |
| 1 (Very pessimistic) | 10.8% (4) | 2.4% (2) | 11.5% (6) | 3.3% (1) | 4.8% (1) | 4.5% (3) | 5.9% (17) |
| 2 | 8.1% (3) | 3.6% (3) | 5.8% (3) | 10.0% (3) | 4.8% (1) | 4.5% (3) | 5.5% (16) |
| 3 | 16.2% (6) | 14.5% (12) | 13.5% (7) | 23.3% (7) | 33.3% (7) | 21.2% (14) | 18.3% (53) |
| 4 | 21.6% (8) | 16.9% (14) | 17.3% (9) | 20.0% (6) | 4.8% (1) | 19.7% (13) | 17.6% (51) |
| 5 | 10.8% (4) | 21.7% (18) | 23.1% (12) | 3.3% (1) | 23.8% (5) | 16.7% (11) | 17.6% (51) |
| 6 (Very optimistic) | 32.4% (12) | 41.0% (34) | 28.8% (15) | 40.0% (12) | 28.6% (6) | 33.3% (22) | 34.9% (101) |
| \*\*Total\*\* | \*\*37\*\* | \*\*83\*\* | \*\*52\*\* | \*\*30\*\* | \*\*21\*\* | \*\*66\*\* | \*\*289\*\* |

**Line 2: Perception vs living time**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P67 | ≤ 1 year | > 1 and ≤ 5 years | > 5 and ≤ 10 years | > 10 and ≤ 15 years | > 15 and ≤ 20 years | > 20 years | Total |
| 1 (Very pessimistic) | 13.5% (5) | 11.0% (8) | 17.8% (8) | 5.3% (1) | 18.2% (2) | 16.1% (10) | 13.8% (34) |
| 2 | 2.7% (1) | 2.7% (2) | 6.7% (3) | 0.0% (0) | 18.2% (2) | 6.5% (4) | 4.9% (12) |
| 3 | 21.6% (8) | 26.0% (19) | 26.7% (12) | 15.8% (3) | 18.2% (2) | 16.1% (10) | 21.9% (54) |
| 4 | 18.9% (7) | 17.8% (13) | 22.2% (10) | 36.8% (7) | 18.2% (2) | 14.5% (9) | 19.4% (48) |
| 5 | 21.6% (8) | 13.7% (10) | 15.6% (7) | 10.5% (2) | 0.0% (0) | 22.6% (14) | 16.6% (41) |
| 6 (Very optimistic) | 21.6% (8) | 28.8% (21) | 11.1% (5) | 31.6% (6) | 27.3% (3) | 24.2% (15) | 23.5% (58) |
| \*\*Total\*\* | 37 | 73 | 45 | 19 | 11 | 62 | 247 |

#### Perception vs Monthly Rent

**Line 1: perception vs monthly rent**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P67 | Prefer not to answer | Less than or equal to \$500,000 | More than \$500,000 and up to \$1,000,000 | More than \$1,000,000 and up to \$1,500,000 | More than \$1,500,000 and up to \$2,000,000 | More than \$2,000,000 and up to \$3,000,000 | Total |
| 1 (Very pessimistic) | 6.2% (1) | 0.0% (0) | 9.3% (7) | 4.5% (1) | 0.0% (0) | 20.0% (1) | 10 |
| 2 | 18.8% (3) | 0.0% (0) | 1.3% (1) | 4.5% (1) | 25.0% (1) | 20.0% (1) | 7 |
| 3 | 12.5% (2) | 15.8% (3) | 14.7% (11) | 18.2% (4) | 0.0% (0) | 20.0% (1) | 21 |
| 4 | 6.2% (1) | 15.8% (3) | 18.7% (14) | 22.7% (5) | 25.0% (1) | 20.0% (1) | 25 |
| 5 | 12.5% (2) | 15.8% (3) | 18.7% (14) | 27.3% (6) | 25.0% (1) | 0.0% (0) | 26 |
| 6 (Very optimistic) | 43.8% (7) | 52.6% (10) | 37.3% (28) | 22.7% (5) | 25.0% (1) | 20.0% (1) | 52 |
| \*\*Total\*\* | \*\*16\*\* | \*\*19\*\* | \*\*75\*\* | \*\*22\*\* | \*\*4\*\* | \*\*5\*\* | \*\*141\*\* |

**Line 2: perception vs monthly rent**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| P67 | Prefer not to answer | ≤ \$500,000 | > \$500,000 – \$1,000,000 | > \$1,000,000 – \$1,500,000 | > \$1,500,000 – \$2,000,000 | > \$2,000,000 – \$3,000,000 | > \$3,000,000 – \$4,000,000 | Total |
| 1 (Very pessimistic) | 23.1% (3) | 11.1% (1) | 13.3% (11) | 8.5% (4) | 8.3% (1) | 28.6% (2) | 0.0% (0) | 22 |
| 2 | 0.0% (0) | 0.0% (0) | 3.6% (3) | 6.4% (3) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 6 |
| 3 | 30.8% (4) | 11.1% (1) | 16.9% (14) | 38.3% (18) | 16.7% (2) | 42.9% (3) | 50.0% (1) | 43 |
| 4 | 0.0% (0) | 11.1% (1) | 22.9% (19) | 21.3% (10) | 16.7% (2) | 14.3% (1) | 50.0% (1) | 34 |
| 5 | 23.1% (3) | 55.6% (5) | 16.9% (14) | 10.6% (5) | 8.3% (1) | 14.3% (1) | 0.0% (0) | 29 |
| 6 (Very optimistic) | 23.1% (3) | 11.1% (1) | 26.5% (22) | 14.9% (7) | 50.0% (6) | 0.0% (0) | 0.0% (0) | 39 |
| \*\*Total\*\* | \*\*13\*\* | \*\*9\*\* | \*\*83\*\* | \*\*47\*\* | \*\*12\*\* | \*\*7\*\* | \*\*2\*\* | \*\*173\*\* |

### Takeaway

#### General summary

Relationship 1 examines how respondents’ overall perception of the Bogotá Metro (P67, an ordinal scale from 1 = “Very pessimistic” to 6 = “Very optimistics ”) varies by three housing characteristics:

- **Ownership (P82)**: Both owners (n=148) and renters (n=141) show a majority of optimistics views (P67=6): 33.1% of owners and 36.9% of renters in Line 1, and 25.7% of owners vs. 22.5% of renters in Line 2 report “Very optimistic.” Pessimistic views (P67=1–2) are uniformly low across both groups

- **Length of Residence (P83)**: Longer-term residents tend to be more optimistics. In Line 1, those living in their home > 5 years account for over 68% of the “Very optimistics” category (P67=6), whereas shorter-term residents (≤ 1 year) make up only 32.4% of that group

- **Monthly Rent (P86)**: Lower-rent tenants report higher optimistic: among those paying ≤ 500 000 COP, 52.6% are “Very optimistics,” compared with 22.7% in the > 1 000 000–1 500 000 COP bracket

#### Suitability for Regression Analysis

- **Variable types & sample size**: As before, P67 is an ordinal outcome; P82 is binary; P83 and P86 are ordinal. Overall N≈289 is adequate, but some P83/P86 cells are small and may need category collapsing.

#### Potential Regression Model

- **Ordinal logistic regression (proportional-odds model)**, with P67 as the outcome and P82 (ownership), P83 (residence duration), and P86 (rent level) as predictors.

- **Sample size and cell counts**: With ~289 respondents for Line 1, there is adequate overall sample size, but several P83 and P86 categories have small cell counts (e.g., only 4 respondents in the “> 1 500 000–2 000 000 COP” rent category).

- **Assumptions**: An ordinal regression model (e.g., proportional odds) requires the proportional-odds assumption; small cells may need collapsing of adjacent categories to ensure stable estimates. After suitable recoding, the data are appropriate for regression.

- **Covariant concerns**: A small sample size combined with limited variation in the predictors across categories reduces statistical power and often leads to non-significant results, making it difficult to detect true effects and increasing the likelihood of failing to reject the null hypothesis.

**Alternative approaches**:

- If the proportional-odd assumption is violated, a multinomial logistic regression on P67.

- If P67 is dichotomized (e.g., optimistic vs. pessimistic), a binary logistic regression.

- For a simpler approximation, treating P67 as continuous and using linear regression, noting the usual caveats about ordinal outcomes. (**Not ideal**)

## Relationship 2: Impact of Metro Project on Housing Rents (increase the sample size by including all household members)

### Hypothesis

- Most residents are renters

- They prefer to live close to mass transit to reduce travel times

### Bivariable table

According to the distribution of housing tenure, approximately 70% of residents living along Bogotá’s Second Metro Line (L2) are renters, significantly higher than those along Line 1 (L1), where the proportion of owners and renters is nearly equal. This supports Hypothesis 1: most residents are renters, particularly in areas affected by the L2 project.

To examine this relationship, we separately analyze the commuting modes of people living near Line 1 and Line 2. For Line 1, the bivariate table is as follows:

**Line 1: Commuting Modes by Tenure**

|  |  |  |
| --- | --- | --- |
| P52\\_group | Owner | Renter |
| BRT | 41 (35.7%) | 42 (36.5%) |
| Bicycle | 8 (7.0%) | 11 (9.6%) |
| Car | 9 (7.8%) | 7 (6.1%) |
| Intermunicipal | 5 (4.3%) | 0 (0.0%) |
| Motorcycle | 11 (9.6%) | 9 (7.8%) |
| Other | 4 (3.5%) | 0 (0.0%) |
| Other Private | 0 (0.0%) | 2 (1.7%) |
| Other Public | 1 (0.9%) | 1 (0.9%) |
| SITP | 12 (10.4%) | 14 (12.2%) |
| Taxi | 6 (5.2%) | 3 (2.6%) |
| Walk | 18 (15.7%) | 26 (22.6%) |

**Line 2: Commuting Modes by Tenure**

|  |  |  |
| --- | --- | --- |
| P52\\_group | Owner | Renter |
| BRT | 22 (41.5%) | 41 (27.9%) |
| Bicycle | 5 (9.4%) | 16 (10.9%) |
| Car | 4 (7.5%) | 9 (6.1%) |
| Motorcycle | 3 (5.7%) | 13 (8.8%) |
| Other | 1 (1.9%) | 5 (3.4%) |
| Other Private | 0 (0.0%) | 1 (0.7%) |
| Other Public | 1 (1.9%) | 2 (1.4%) |
| SITP | 8 (15.1%) | 18 (12.2%) |
| Taxi | 1 (1.9%) | 2 (1.4%) |
| Walk | 8 (15.1%) | 40 (27.2%) |

In terms of travel mode, renters along Line 1 use BRT (36.5%) and SITP (12.2%) at slightly higher rates than owners, indicating a relatively stronger reliance on mass transit. However, along Line 2, despite renters making up a larger share, they use BRT (27.9%) and SITP (12.2%) less than owners (41.5% and 15.1%, respectively), while the proportion of walking increases significantly (27.2%). This suggests that renters in the L2 area tend to live closer to their workplaces or metro stations, thereby reducing travel time, which supports the second hypothesis that they prefer to live near mass transit to shorten commute durations.

## Relationship 3: Income vs Willingness to Pay

### Research Question

**Household income vs Willingness to pay**

The survey questions involve `P50` and `P81`

### Hypothesis:

The higher the income is, the higher the amount they are willing to pay for the Metro fare

### Bivariate table

**Line 1: Income vs Willingness to pay**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| P50 | $0 - $3,200 | $3,200 - $3,500 | $3,500 - $3,800 | $3,800 - $4,100 | $4,100 - $4,500 | More than $4,500 | Total |
| 0 - $400,000 | 2.4% (5) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 1.7% (5) |
| $400,001 - $800,000 | 6.7% (14) | 11.1% (6) | 11.1% (2) | 0.0% (0) | 50.0% (1) | 0.0% (0) | 8.0% (23) |
| $800,001 - $1,160,000 | 9.6% (20) | 9.3% (5) | 5.6% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 9.0% (26) |
| $1,160,001 - $1,500,000 | 9.6% (20) | 11.1% (6) | 22.2% (4) | 20.0% (1) | 0.0% (0) | 0.0% (0) | 10.7% (31) |
| $1,500,000 - $2,000,000 | 5.7% (12) | 13.0% (7) | 16.7% (3) | 40.0% (2) | 0.0% (0) | 0.0% (0) | 8.3% (24) |
| $2,000,001 - $2,500,000 | 8.6% (18) | 11.1% (6) | 5.6% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 8.7% (25) |
| $2,500,001 - $3,500,000 | 9.6% (20) | 16.7% (9) | 11.1% (2) | 20.0% (1) | 0.0% (0) | 100.0% (1) | 11.4% (33) |
| $3,500,001 - $4,900,000 | 6.7% (14) | 5.6% (3) | 5.6% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 6.2% (18) |
| $4,900,001 - $6,800,000 | 2.4% (5) | 5.6% (3) | 11.1% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 3.5% (10) |
| $6,800,001 - $9,000,000 | 1.9% (4) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 1.4% (4) |
| Over $9,000,000 | 0.0% (0) | 0.0% (0) | 0.0% (0) | 20.0% (1) | 0.0% (0) | 0.0% (0) | 0.3% (1) |
| not reported | 36.8% (77) | 16.7% (9) | 11.1% (2) | 0.0% (0) | 50.0% (1) | 100.0% (1) | 30.8% (89) |
| \*\*Total\*\* | \*\*100.0% (209)\*\* | \*\*100.0% (54)\*\* | \*\*100.0% (18)\*\* | \*\*100.0% (5)\*\* | \*\*100.0% (2)\*\* | \*\*100.0% (1)\*\* | \*\*100.0% (289)\*\* |

**Line 2: Income vs Willingness to pay**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| P50 | $0 - $3,200 | $3,200 - $3,500 | $3,500 - $3,800 | $3,800 - $4,100 | More than $4,500 | Total |
| 0 - $400,000 | 3.8% (7) | 0.0% (0) | 7.1% (1) | 0.0% (0) | 0.0% (0) | 3.2% (8) |
| $400,001 - $800,000 | 7.0% (13) | 4.9% (2) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 6.1% (15) |
| $800,001 - $1,160,000 | 12.9% (24) | 4.9% (2) | 7.1% (1) | 0.0% (0) | 0.0% (0) | 10.9% (27) |
| $1,160,001 - $1,500,000 | 22.0% (41) | 19.5% (8) | 28.6% (4) | 0.0% (0) | 33.3% (1) | 21.9% (54) |
| $1,500,000 - $2,000,000 | 15.1% (28) | 24.4% (10) | 7.1% (1) | 66.7% (2) | 33.3% (1) | 17.0% (42) |
| $2,000,001 - $2,500,000 | 11.3% (21) | 9.8% (4) | 7.1% (1) | 0.0% (0) | 33.3% (1) | 10.9% (27) |
| $2,500,001 - $3,500,000 | 6.5% (12) | 14.6% (6) | 14.3% (2) | 0.0% (0) | 0.0% (0) | 8.1% (20) |
| $3,500,001 - $4,900,000 | 9.7% (18) | 7.3% (3) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 8.5% (21) |
| $4,900,001 - $6,800,000 | 2.2% (4) | 7.3% (3) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 2.8% (7) |
| $6,800,001 - $9,000,000 | 0.0% (0) | 0.0% (0) | 7.1% (1) | 0.0% (0) | 0.0% (0) | 0.4% (1) |
| Over $9,000,000 | 0.5% (1) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.0% (0) | 0.4% (1) |
| Not Reported | 9.1% (17) | 7.3% (3) | 21.4% (3) | 33.3% (1) | 0.0% (0) | 9.7% (24) |
| \*\*Total\*\* | \*\*100.0% (186)\*\* | \*\*100.0% (41)\*\* | \*\*100.0% (14)\*\* | \*\*100.0% (3)\*\* | \*\*100.0% (3)\*\* | \*\*100.0% (247)\*\* |

### Key Takeaway

#### General summary

**In summary, a significant majority portion of the residents only willingness to pay price lower than $3,200. (75% for Line 2 and 72% for Line 1)**

- **Positive association**: Both for Line 1 and Line 2, mid- to upper-middle income households (especially those earning between COP 1.16 million and COP 2 million) make up progressively larger shares of respondents in the higher willingness-to-pay categories (≥ COP 3 800), while lower-income groups and non-reporters dominate the lowest category (≤ COP 3 200). For example, in Line 1 households earning COP 1.16–1.5 million account for 22.2% of those willing to pay COP 3 500–3 800, compared with just 9.6% of the lowest-income group; meanwhile, 36.8% of the lowest-willingness group did not report income :contentReference[oaicite:0]{index=0}.

- **Non-response and sparse extremes**: Income non-response is substantial in Line 1 (30.8% in the lowest willingness category) and present in Line 2 (9.7%), and the highest income brackets (above COP 4.9 million) and highest willingness categories each contain very few observations (often fewer than 5), which reduces reliability at the tails :contentReference[oaicite:1]{index=1}.

#### Suitability for Regression Analysis

- **Outcome**: Willingness to pay (P81) is inherently ordinal (six ordered categories).

- **Predictor**: Income group (P50) is also ordinal but includes a large “Not reported” category, which may need to filter out

- **Sample size**: N≈289 for Line 1 and N≈247 for Line 2, adequate overall but with very small cell counts in some combinations (e.g., highest income × highest willingness). These sparse cells will need category collapsing or exclusion to avoid unstable estimates. These may lead to yield bias regression results.

#### Potential Regression Model

- **Ordinal logistic regression** (proportional-odds model): Use P81 (willingness) as the outcome and P50 (income bracket) as the predictor.

- **Category consolidation**: Merge extreme income brackets (e.g., all > COP 2 million) into one or two top categories and combine the lowest willingness categories (e.g., ≤ COP 3 200 vs > COP 3 200) if needed to ensure minimal cell counts.

- **Assumption checking**: Test the proportional-odds assumption (e.g., Brant test). If violated, consider a partial proportional odds model or switch to multinomial logistic regression.

#### Alternative approaches

- **Binary logistic regression**: Dichotomize willingness (e.g., low vs high) and model as a binary outcome.

- **Linear regression proxy**: Assign mid-points to P81 and P50 and treat them as continuous in a linear model—acceptable as a rough check, but note the ordinal nature.

## Relationship 4: Willingness to use vs transportation for work

### Research Question

**Willingness to use vs transportation mode for work**

The survey question involved `P52`, `P78`

### Hypothesis

The possibility of a mode shift based on the transportation modes

### Bivariate table

**Line 1: Willingness to use vs transportation for work**

|  |  |  |  |
| --- | --- | --- | --- |
| P52 | Yes | No | Total |
| Bicycle | 6.7% (18) | 4.8% (1) | 6.6% (19) |
| BRT | 29.1% (78) | 23.8% (5) | 28.7% (83) |
| Car | 5.2% (14) | 9.5% (2) | 5.5% (16) |
| Intermunicipal | 1.5% (4) | 4.8% (1) | 1.7% (5) |
| Motorcycle | 6.3% (17) | 14.3% (3) | 6.9% (20) |
| Other | 1.5% (4) | 0.0% (0) | 1.4% (4) |
| Other Private | 0.4% (1) | 4.8% (1) | 0.7% (2) |
| Other Public | 0.7% (2) | 0.0% (0) | 0.7% (2) |
| SITP | 9.0% (24) | 9.5% (2) | 9.0% (26) |
| Taxi | 3.0% (8) | 4.8% (1) | 3.1% (9) |
| Walk | 16.0% (43) | 4.8% (1) | 15.2% (44) |
| Not reported or unemployed | 20.5% (55) | 19.0% (4) | 20.4% (59) |
| \*\*Total\*\* | \*\*100.0% (268)\*\* | \*\*100.0% (21)\*\* | \*\*100.0% (289)\*\* |

**Line 2: Willingness to use vs transportation for work**

|  |  |  |  |
| --- | --- | --- | --- |
| P52 | Yes | no | Total |
| Bicycle | 8.0% (18) | 13.6% (3) | 8.5% (21) |
| BRT | 27.1% (61) | 9.1% (2) | 25.5% (63) |
| Car | 4.9% (11) | 9.1% (2) | 5.3% (13) |
| Motorcycle | 5.8% (13) | 13.6% (3) | 6.5% (16) |
| Other | 2.7% (6) | 0.0% (0) | 2.4% (6) |
| Other Private | 0.0% (0) | 4.5% (1) | 0.4% (1) |
| Other Public | 0.9% (2) | 4.5% (1) | 1.2% (3) |
| SITP | 11.1% (25) | 4.5% (1) | 10.5% (26) |
| Taxi | 1.3% (3) | 0.0% (0) | 1.2% (3) |
| Walk | 20.0% (45) | 13.6% (3) | 19.4% (48) |
| Not Reported or unemployed | 18.2% (41) | 27.3% (6) | 19.0% (47) |
| \*\*Total\*\* | \*\*100.0% (225)\*\* | \*\*100.0% (22)\*\* | \*\*100.0% (247)\*\* |

### Key Takeaway

Overall, since there are only a low share of the sample answers `no`(7% for line 1, and 9% for line 2), the percentage tend to bump up due to random effect.

- **Strongest adoption potential among transit and pedestrian commuters**: BRT users (29.1% for Line 1; 27.1% for Line 2), walk commuters (16.0%; 20.0%) and SITP riders (9.0%; 11.1%) make up the largest shares of those willing to use the Metro.

- **Lower willingness among private and active-mode users**: Car commuters show modest willingness (5.2%; 4.9%) but higher unwillingness (9.5%; 9.1%), motorcycle users have moderate willingness (6.3%; 5.8%) yet substantial reluctance (14.3%; 13.6%), and bicycle users exhibit lower willingness (6.7%; 8.0%) alongside non-negligible unwillingness (4.8%; 13.6%).

- **Mixed attitudes in unemployed/non-reported group**: This segment accounts for 20.5% (Line 1) and 18.2% (Line 2) of the willing respondents, but rises to 19.0% and 27.3% among the unwilling, indicating heterogeneity in this cohort.

- **Consistent patterns across Lines 1 and 2**: The distribution of willingness by current work mode remains stable for both metro lines, underscoring mode-specific affinities toward adoption.

#### Suitability for Regression Analysis

- **Outcome**: Willingness to use Metro (Binary: Yes or No).

- **Predictor**: Current commute mode (nominal: BRT, walk, SITP, car, motorcycle, bicycle, unemployed/non-reported).

- **Sample size**: N≈289 for Line 1 and N≈247 for Line 2; overall adequate but some due to the lower sample size of the response `no`, the acutual covariant or useful sample size may not be adequate.

#### Potential Regression Model

- **logistic regression**: Model the three-level willingness outcome with commute mode as predictor.