

201B First Part Final Assignment

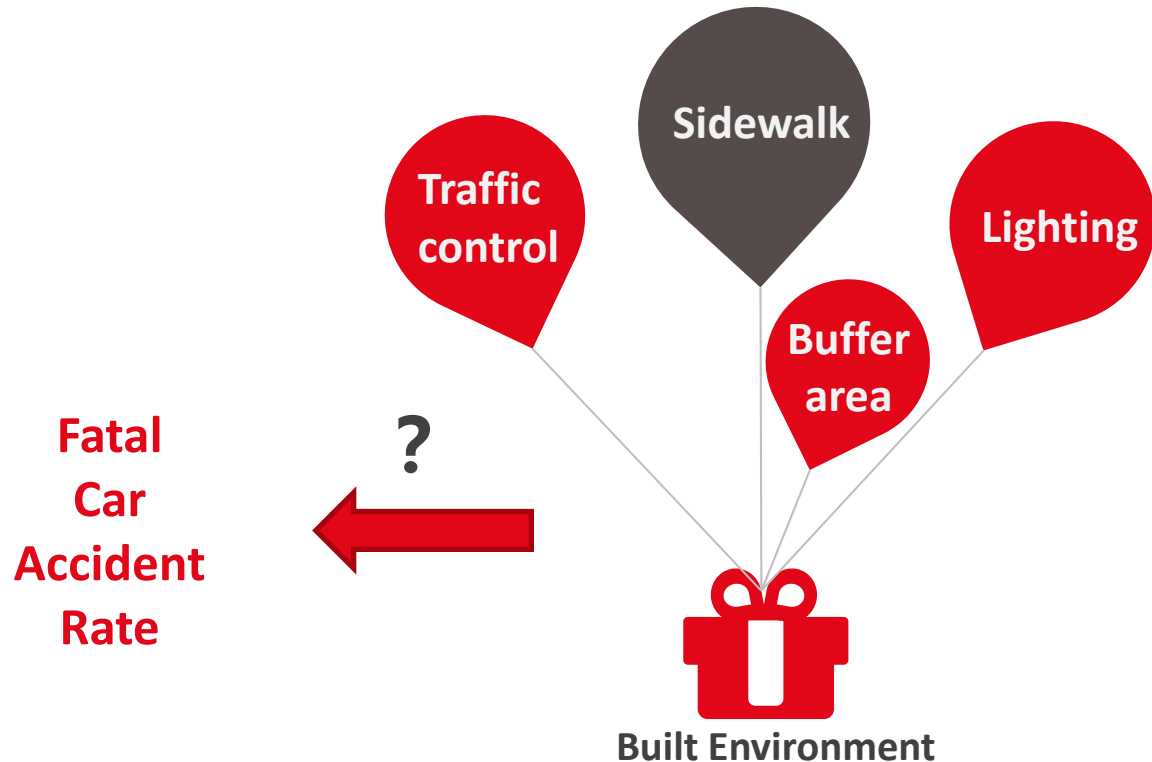
What factors are related to the fatal car accident rate in the neighborhood?

Yuquan Zhou 2020.3.2

01 Introduction and Relevance

Nearly **1.25 million** people die in road crashes, which ranks as the **9th** leading cause of death globally each year, making car accident a **major threat to public health**^[1]. Therefore, I am motivated to explore what planning factors are related to the car accident rate. Whether there are things that we as planners can do to decrease the car accident rate.

There is emerging evidence on the association between the built environment and travel behavior^[2]. Based on that, the study makes the hypothesis that build environment is associated with car accident rate and explores what aspects of built environment and planning aspect are associated with fatal car accident rate and how.



Source:

[1] <https://www.asirt.org/safe-travel/road-safety-facts/>

[2] Estupiñán, N., & Rodríguez, D. A. (2008). The relationship between urban form and station boardings for Bogota's BRT. Transportation Research Part A: Policy and Practice, 42(2), 296-306.

Source of data

Pedestrian count data in Bogota collected in 2005

Variable type: Segment level N= 338

Neighborhood level N= 71

Since the fatal car accident data is in neighborhood level, the study aggregated segment level data into neighborhood level by the neighborhood it belongs to.

Original segment level data

Panhandle

Presence of homeless person on segment (1=yes)

Average value in each neighborhood

Percent of segment with panhandle presence in the neighborhood

Way finding aids

Presence of wayfinding aids on segment (1=yes)

Average value in each neighborhood

Percent of segment with wayfinding aids in the neighborhood

Buffer width

Buffer width between sidewalk and road (0= no buffer, 1= <2 meters; 2= 2 meters and more)

Recode into dummy variables:
0=buffer <2 meters; 1 = buffer>2 meters
Calculate average in each neighborhood

Percent of segment with buffer width >2 meters in the neighborhood

Traffic Control

Percent of all potential pedestrian-related traffic control elements available on segment

Average value in each neighborhood

Average percentage of all pedestrian-related traffic control elements in the neighborhood

Sidewalk Width

(1= <2 meters; 2=2-4 meters; 3= > 4 meters)

Recode into dummy variables:
0=width <4 meters; 1 =width >=4meters
Calculate average in each neighborhood

Percent of segment with sidewalk width>4 meters in the neighborhood

Source of data

Pedestrian count data in Bogota collected in 2005

Universe: neighborhood N=71

Dependent Variable

Car accidents rate:

Car accidents that reported at least one fatality per thousand residents

Key Independent Variable

1. Wayfinding aids:

Percent of segment with wayfinding aids in the neighborhood

2. Buffer width

Percent of segment with buffer width between sidewalk and road wider than 2 meters in the neighborhood

3. Traffic Control

Average percentage of all potential pedestrian-related traffic control elements available of segment in neighborhood (stoplight + pedestrian signal + stop sign + crosswalk+ overpass)

4. Sidewalk Width

Percent of segment with sidewalk width wider than 4 meters in the neighborhood

Control Variables

1. SES status

Socio-economic stratum of neighborhood (1-6; 6 being highest)

2. Road Density

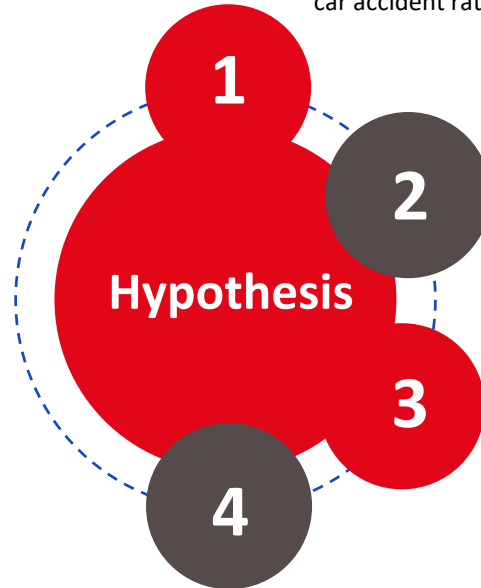
Density of roads (mi/sq mi) in neighborhood

3. Panhandle

Percent of segment with panhandle presence in the neighborhood

Regression Type: Multi-linear regression

The study adopts **Multivariable linear regression** to explore the relationship between fatal car accident rate and the key independent variables listed on the right.



Wayfinding aids

The percentage of segment with wayfinding aids in the neighborhood is **strongly negatively** associated with car accident rate.

Buffer width

The percentage of segment with buffer width between sidewalk and road wider than 2 meters in the neighborhood is **strongly negatively** associated with car accident rate.

Traffic Control

The average percentage of all potential pedestrian-related traffic control elements available of segment in neighborhood is **strongly negatively** associated with car accident rate.

Sidewalk Width

The percentage of segment with sidewalk width wider than 4 meters in the neighborhood is **strongly negatively** associated with car accident rate.

1. Wayfinding aids

Table 1: Data description of percent of segment with wayfinding aids in the neighborhood

Count	Mean	Std	Min	25%	50%	75%	Max
71	34.91%	22.50%	0.00%	20%	33.33%	50.00%	100.00%

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota

Data source: Pedestrian count data in Bogota (2005)

2. Buffer width

Table 2: Data description of percent of segment with buffer width >2 meters in the neighborhood

Count	Mean	Std	Min	25%	50%	75%	Max
71	5.42%	12.27%	0.00%	0.00%	0.00%	0.00%	60.00%

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota

Data source: Pedestrian count data in Bogota (2005)

3. Traffic Control

Table 3: Data description of average percentage of traffic control elements of segment in neighborhood

Count	Mean	Std	Min	25%	50%	75%	Max
71	40.25%	15.48%	12.00%	28.00%	40.00%	55.00%	72.00%

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota

Data source: Pedestrian count data in Bogota (2005)

4. Sidewalk Width

Table 1: Data description of percent of segment with sidewalk width > 4 meters in the neighborhood

Count	Mean	Std	Min	25%	50%	75%	Max
71	19.20%	17.30%	0.00%	0.00%	20.00%	29.17%	60.00%

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota

Data source: Pedestrian count data in Bogota (2005)

On average, 34.91%, 5.42%, 34.91% of segment in the neighborhood has wayfinding aids, buffer wider than 2 meters, and sidewalk more than 4 meters, respectively, and mean of average percentage of traffic control elements in neighborhood is 40.25. Especially, only a few neighborhoods has buffer wider than 2 meters, and more than 75% of the neighborhood doesn't have buffer wider than 2 meters.

Dependent Variable: Car accidents rate:

Table 5: Data description of car accidents that reported at least one fatality per thousand

Count	Mean	Std	Min	25%	50%	75%	Max
71	11.85%	5.34%	4.30%	7.50%	11.20%	15.40%	28.54%

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota
Data source: Pedestrian count data in Bogota (2005)

On average, the fatal car accident rate in the neighborhood is 11.85% per thousand.

Control Variables

1. SES status

Table 6: Data description of Socio-economic stratum of neighborhood

Count	Mean	Std	Min	25%	50%	75%	Max
71	3.1173	0.4960	2.290	2.720	3.000	3.590	4.370

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota
Data source: Pedestrian count data in Bogota (2005)

On average, 42.68% of segments in the neighborhood have panhandles presence. The average SES status among 71 neighborhoods is 3.1173, and the average road density among 71 neighborhoods is 9.534 mi/mi².

2. Road Density

Table 7: Data description of road density (mi/mi²) of neighborhood

Count	Mean	Std	Min	25%	50%	75%	Max
71	9.534	2.291	4.870	7.830	9.770	10.830	16.320

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota
Data source: Pedestrian count data in Bogota (2005)

3. Homeless people

Table 8: Data description of percent of segment with homeless person presence in the neighborhood

Count	Mean	Std	Min	25%	50%	75%	Max
71	42.68%	34.21%	0.00%	20.00%	40.00%	60.00%	100.00%

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota
Data source: Pedestrian count data in Bogota (2005)

02 Data and Method

Bivariate relationship

Figure 1: Bivariate relationship between **wayfinding aids** and car accident rate

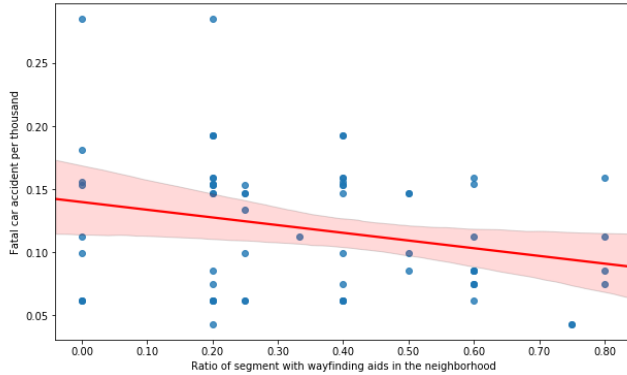


Figure 2: Bivariate relationship between **buffer width** and car accident rate

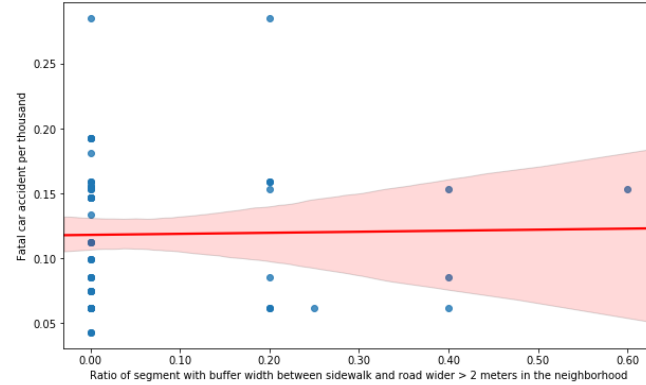


Figure 3: Bivariate relationship between **traffic control** and car accident rate

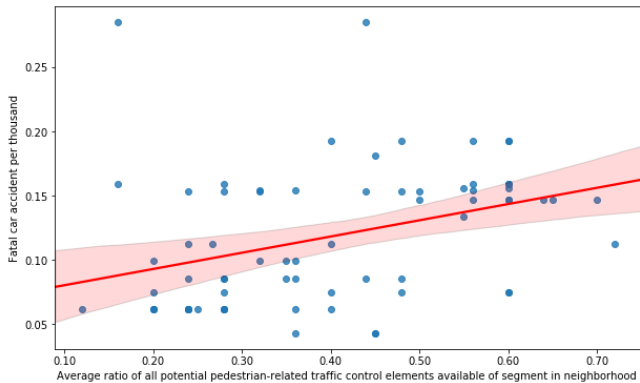
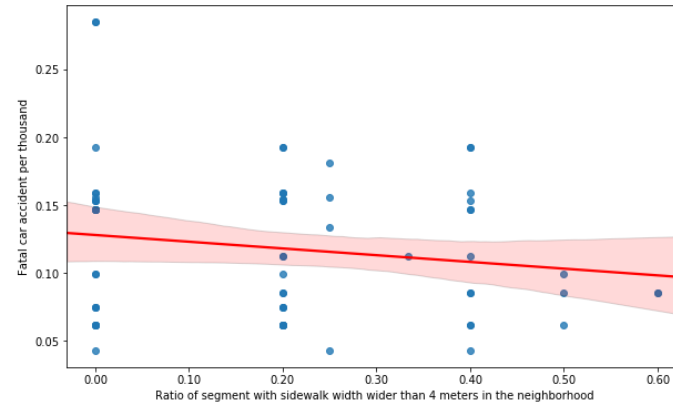


Figure 4: Bivariate relationship between **sidewalk width** and car accident rate



*N=71, Universe:
all researched
neighborhood
around bus rapid
transit stops in
Bogota
Data source:
Pedestrian count
data in Bogota
(2005)

The independent variables explain 48.3% of the dependent variable (Car accident rate per thousand)

Table 9: Overall Model Fit

Dependent Variable	No. Observations:	R-squared:	Adj. R-squared:	F-statistic
Car Accident Rate per thousand	71	0.483	0.426	8.414

Table 10: Coefficient and other statistics of each independent variable





variables	Coefficient	Standard error	t	P> t	[0.025	0.975]
Wayfinding aids	-0.043	0.0232	-1.8615	0.0673*	-0.0895	0.0032
Buffer width	0.1132	0.0440	2.5722	0.0125**	0.0253	0.2012
Traffic Control	0.0608	0.0426	1.4269	0.1585	-0.0243	0.1459
Sidewalk Width	-0.0430	0.0302	-1.4257	0.1589	-0.1033	0.0173
SES Status	-0.0285	0.0105	-2.7142	0.0086***	-0.0495	-0.0075
Road density	-0.0047	0.0022	-2.1398	0.0363**	-0.0092	-0.0003
panhandle	0.0824	0.0186	4.4318	0.0000***	0.0452	0.1195
Intercept	0.2100	0.0429	4.8932	0.0000	0.1243	0.2958




*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota
Data source: Pedestrian count data in Bogota (2005)

Fatal car accident rate per thousand = $-0.042 * \text{Way finding aids} + 0.1132 * \text{buffer width} + 0.0608 * \text{Traffic Control} - 0.0430 * \text{Sidewalk Width} - 0.0285 * \text{SES Status} - 0.0047 * \text{Road Density} + 0.0824 * \text{Panhandle} + 0.2100$

WHILE HOLDING OTHER VARIABLES CONSTANT...

Car accidents that reported at least one fatality per thousand

-  **Decreases 4.32%** when percentage of segment with **wayfinding aids** in the neighborhood increases from 0% to 100%.
-  **Increases 11.32%** when the percentage of segment with **buffer width** between sidewalk and road wider than 2 meters in the neighborhood increases from 0% to 100%.
-  **Increases 6.08%** when average percentage of **traffic control** elements available of segment in neighborhood increases from 0% to 100%.
-  **Decreases 4.30%** when percentage segment with **sidewalk width** wider than 4 meters in the neighborhood increases from 0% to 100%.

-  **Decreases 2.85%** when socio-economic stratum of neighborhood (**SES**) increases by 1 unit.
-  **Decreases 0.47%** when the road density increases 1 mi/mi² in the neighborhood increases
-  **Increases 8.24%** when average percentage of segment with homeless person presence in neighborhood increases from 0% to 100%.

03 Results

Multicollinearity checks

Table 11: Correlation between independent variables

variables	Signposts	Buffer width	Sidewalk Width	SES Status	Road density	panhandle
Wayfinding aids	1.0000	-0.0214	0.0504	0.0963	-0.2093	0.1446
Buffer width	-0.0214	1.0000	-0.3742	-0.0130	0.1985	-0.1610
Traffic Control	0.0504	-0.3742	1.0000	-0.2190	0.0210	0.1911
Sidewalk Width	0.0963	-0.0130	-0.2190	1.0000	-0.2486	0.0879
SES Status	-0.2093	0.1985	0.0210	-0.2486	1.0000	-0.0862
Road Density	0.1446	-0.1610	0.1911	0.0879	-0.0862	1.0000
panhandle	-0.1724	-0.3065	0.5882	-0.0183	-0.0046	0.2020

All the R value of correlation between two independent variable is smaller than 0.8.

Table 12: Variance inflation factor (VIF) of each independent variable

variables	Signposts	Buffer width	Sidewalk Width	SES Status	Road density	panhandle
VIF	1.1639	1.2476	1.1649	1.1589	1.1005	1.7290

All the VIF value is smaller than 4.0

*N=71, Universe: all researched neighborhood around bus rapid transit stops in Bogota
Data source: Pedestrian count data in Bogota (2005)

There is **no multicollinearity** in this regression.

Figure 5: Relationship between residuals and predicted values

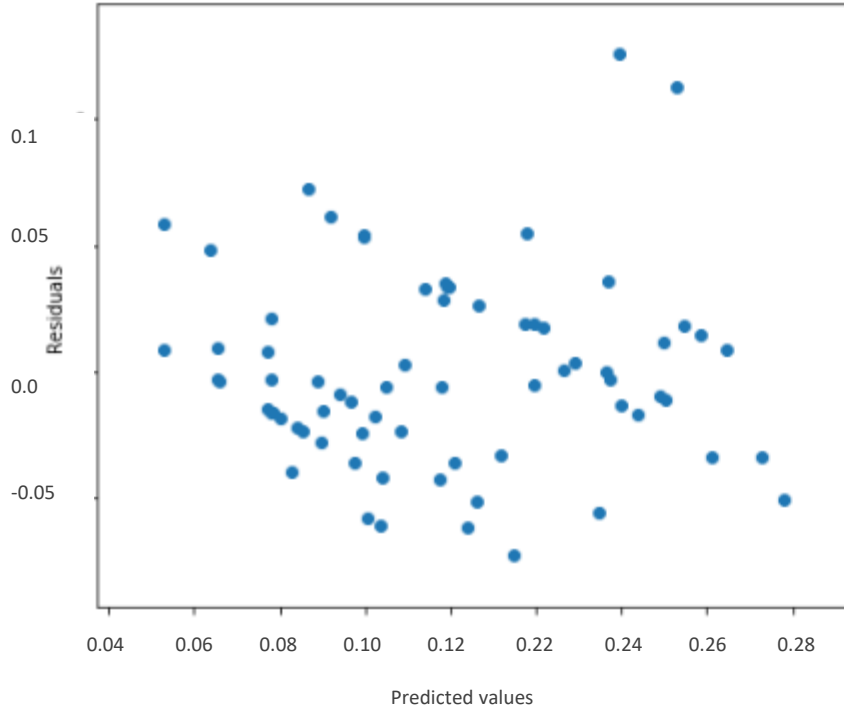
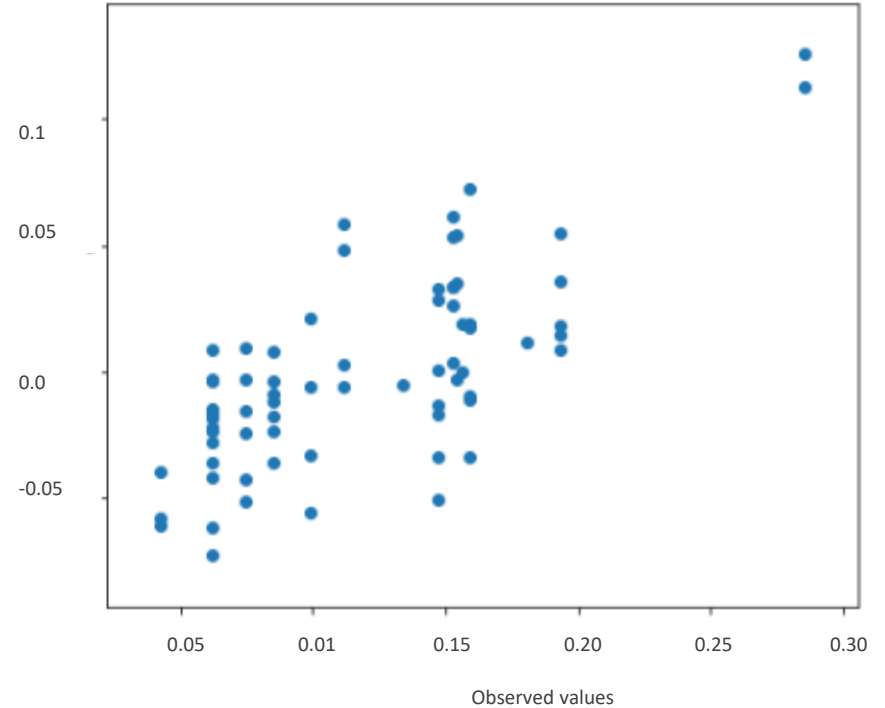


Figure 6: Relationship between residuals and observed values



The plot looks normal and there is **no heteroscedasticity** in this regression.

Car accidents that reported at least one fatality per thousand is

Positive
Associates
with

Fatal car accident rate per thousand increases

Buffer width**

11.32% when the percentage of segment with buffer between sidewalk and road >2 meters in the neighborhood increases from 0% to 100%.

Traffic Control (not significant)

6.08% when average percentage of traffic control elements available of segment in neighborhood increases from 0% to 100%.

Panhandle***

8.24% when average percentage of segment with panhandle presence in neighborhood increases from 0% to 100%

Fatal car accident rate per thousand decreases

Wayfinding aids*

4.32% when percentage of segment with wayfinding aids in the neighborhood increases from 0% to 100%.

Sidewalk Width (not significant)

4.30% when percentage segment with sidewalk width > 4 meters in the neighborhood increases from 0% to 100%.

SES Status***

2.85% when socio-economic stratum of neighborhood (SES) increases by 1 unit.

Road density**

0.47% when the road density increases 1 mi/mi² in the neighborhood.

Negative
Associates
with

Wayfinding aids

The result is consistent with the hypothesis:

The percentage of segment with wayfinding aids in the neighborhood is **negatively** associated with car accident rate at 90% confidence level.

Why:

The more comprehensive the wayfinding aids in the neighborhood, the less likely that the pedestrians wander around and try to figure out which direction to go, which distracts their attention and increases the risk of being hit by a car.

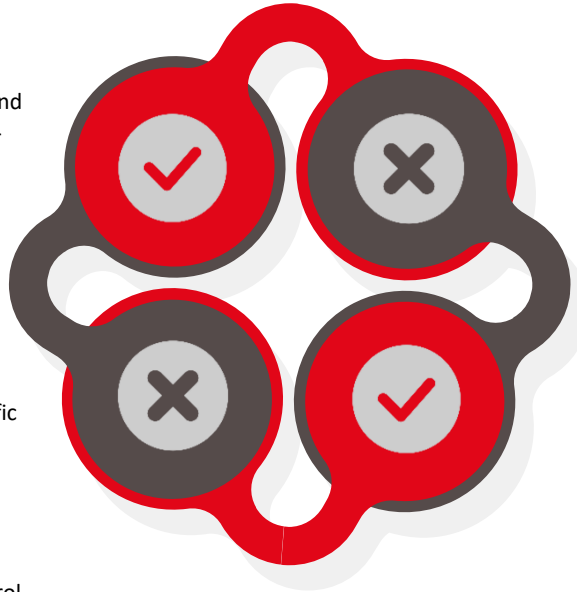
Traffic Control

The result contradicts the hypothesis but not significant:

The average percentage of all potential pedestrian-related traffic control elements available of segment in neighborhood is **positively** associated with car accident rate but not statistically significant.

Why:

A possible explanation is that all pedestrian related traffic control listed in the dataset serves for the same purpose of preventing vehicle collisions that involve a bike or pedestrian. However, the higher percentage the traffic controls does not necessarily contribute to better prevention but leads to more confusion.



Buffer width

The result contradicts the hypothesis:

The percentage of segment with buffer > 2 meters in the neighborhood is **positively** associated with car accident rate at 95% confidence level

Why:

One possible reason is the when the buffer width > 2 meters, homeless people may sleep, panhandles may beg, and even regular pedestrians may rest and chat on the buffer area, which increases their risk of being hit by cars.

Sidewalk Width

The result is consistent with the hypothesis but not significant

The percentage of segment with sidewalk width wider than 4 meters in the neighborhood is strongly **negatively** associated with car accident rate.

Why:

The wider the sidewalk means the more buffer area pedestrians could dodge when a collision happens. However, it is not a relevant factor when collision between vehicles or high-speed severe collision happens.

01



SES Status

Socio-economic stratum of neighborhood is **negatively** associated with car accident rate at 99% confidence level.

Why:

The higher SES Status indicates the better transportation related infrastructure and service which may contribute to lower fatal car accident rate.

02



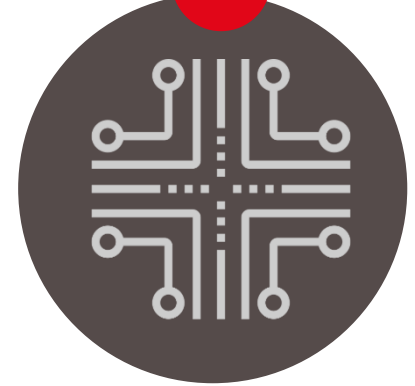
Panhandle

Percent of segment with panhandle presence in the neighborhood is **positively** associated with car accident rate at 95% confidence level.

Why:

One explanation is that the vehicle collision that results in the death of panhandles makes up a significant proportion of the fatal car accident rate. Another possible explanation is that the panhandler distract drivers' attention or block drivers' vision which may result in fatal car accident.

03



Road Density

Density of roads in neighborhood is **negatively** associated with car accident rate at 95% confidence level.

Why:

The higher road density in the neighborhood may lead to slower speed the vehicles travel, which gives drivers more time to react when there is emergency on the road and results in fewer fatal car accidents.

Other variables

◆ Lighting

There is lighting variable in the existing dataset, but the regression outcome is hardly significant. A possible reason is a large majority of the neighborhood in the dataset have 100% lighting on the road.

◆ Speed bump

The variable is not in the dataset but may be an influencing factor since vehicle speed is highly related to fatal car accident.

◆ Bike Lanes

The variable is not in the dataset but may reflect the collision between bicyclist and vehicles or pedestrians.

Recommendation



Improve wayfinding aids on the sidewalk

The study result shows that way finding aids is helpful in reducing the fatal car accident rate.



Design buffer between sidewalk and road inaccessible to pedestrians

Buffer between sidewalk and road is supposed to keep pedestrians from the road and planning the buffer inaccessible to pedestrians may better ensure their safety when collision happened on the road.



Guide the panhandles to safer place away from the road

The study shows that the panhandles are significantly related to fatal car accident rate in the neighborhood. Therefore, guiding the panhandles safer place away from the road is good for their own safety and a way to keep drivers from being distracted by panhandles.