

Data 604 - Final Project - Appendix L

Thanusha Balasubramanian, Elise Beaupre

2025-12-03

Data 604 - Final Project - Linear Regression

```
data <- read_csv("Appendix_K_Overall_Comparison.csv")
```

```
## Rows: 204 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (3): comm_code, comm_name, assessment_class
## dbl (8): median_assessed_value, service_count, service_rank, total_crime_202...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(data)
```

```
## # A tibble: 6 x 11
##   comm_code comm_name      median_assessed_value assessment_class service_count
##   <chr>      <chr>              <dbl> <chr>              <dbl>
## 1 SPR       SPRUCE CLIFF             362000 Low              2
## 2 RUT       RUTLAND PARK             398500 Low              1
## 3 RRC       RICARDO RANCH             73000 Low              0
## 4 LEB       LEWISBURG                146750 Low              0
## 5 KSH       KEYSTONE HILLS           170000 Low              0
## 6 PAT       PATTERSON                 510500 Low              1
## # i 6 more variables: service_rank <dbl>, total_crime_2024 <dbl>,
## #   safety_rank <dbl>, affordability_rank <dbl>, overall_score <dbl>,
## #   overall_rank <dbl>
```

```
model = lm(median_assessed_value~service_count + total_crime_2024
, data = data)
summary(model)
```

```
##
## Call:
## lm(formula = median_assessed_value ~ service_count + total_crime_2024,
##     data = data)
##
## Residuals:
```

```

##      Min      1Q   Median      3Q      Max
## -1265735 -500291 -271796      37 41554781
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1338735     330572   4.050 7.32e-05 ***
## service_count     -76036     318658  -0.239   0.812
## total_crime_2024  -4788       4085  -1.172   0.243
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3006000 on 201 degrees of freedom
## Multiple R-squared:  0.01292,    Adjusted R-squared:  0.003097
## F-statistic: 1.315 on 2 and 201 DF,  p-value: 0.2707

```

The p-value of the model overall is over $\alpha = 0.05$ so overall the difference in median assessed value by community can not be explained by a linear model with only the crime rates, and the amount of amenities.

Additionally, neither `service_count` or `total_crime_2024` have p-values over 0.05 indicating neither variables are significant predictors of the median assessed value of residential property in Calgary.