To: The Manager (314 words)

From: Sashank Addanki Venkata, Rithik Rathinavel Ragupathi, Akshay Reddy Kaidapuram, Rithik Rathinavel Ragupathi, Irfan Saleemudeen Kalidindi, Saketh Varma.

Subject: Exploratory Data Analysis of Denver's Non-airline Revenues

Date: 11 October 2024

Summarize non-airline revenues from year to year and analyze the trend. Also, forecast non-airline revenues to produce monthly revenue for 2022

**Major Finding(s)**

Drivers of Revenue: Regression analysis suggests that major contributors to concession revenue include parking, rental cars, and ground transportation; among these three, the greatest is parking as shown in Exhibit A.

Model Performance: From the model output result, the model has achieved an R-square of 87.2% (Exhibit B), which means that a large part of the concession revenue variance is explained by the variables. These refinements eliminate problems of multicollinearity, non-normality, and heteroscedasticity.

Improvements to Regression: Feature engineering-interaction and polynomial terms, regularization techniques-Lasso, and Ridge, log transformations(Exhibit C).

Violations resolved: Multicollinearity and non-normal residuals were reduced by VIF analysis, as shown in Exhibit D, and data transformation to improve the reliability of the model.

**Recommendations**

Revenue Optimization: Allocate more marketing resources to streams that make the most impact, such as parking and rental cars. Use dynamic pricing based on historical trends to realize revenue.

Operational Efficiency: Putting to better use the resources available for parking and renting services. Automation of repetitive tasks, such as ticketing and car parking management, would be enabled through mobile applications, which would also facilitate the booking process for customers.

Data-Driven Decision Making: Leverage predictive analytics to forecast demand about parking and rental car services. Based on current usage, through real-time monitoring-e.g., via parking sensors-dynamically observe and change operations.

Staff Optimization: Provide a staffing forecast based on historic data; cross-train employees for flexibility to capture the right staffing during peak and off-peak hours.

**Analytical Overview**

Firstly the data had missing values in the ‘Month’ column which we replaced with the relevant data using the previous data as a reference. Then we performed exploratory data analysis on the revenue columns through which we were able to plot out some relationships and get a better understanding of where the focus is to be concentrated.

**Appendix**

Exhibit A

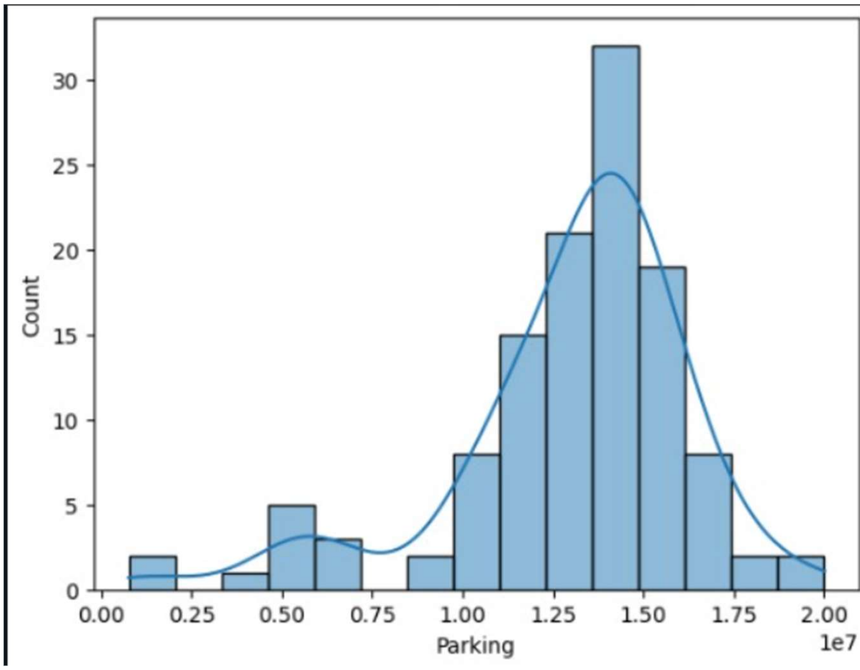

Exhibit B

```
                            OLS Regression Results
==============================================================================
Dep. Variable:             Concession   R-squared:                       0.872
Model:                            OLS   Adj. R-squared:                  0.869
Method:                 Least Squares   F-statistic:                     263.8
Date:                Fri, 11 Oct 2024   Prob (F-statistic):           1.23e-51
Time:                        22:48:28   Log-Likelihood:                -1743.8
No. Observations:                 120   AIC:                             3496.
Df Residuals:                     116   BIC:                             3507.
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       7.833e+05   1.96e+05      4.001      0.000    3.95e+05    1.17e+06
Parking        0.1070      0.019      5.649      0.000       0.070       0.145
Rental Car     0.1906      0.044      4.321      0.000       0.103       0.278
Ground         1.9681      0.160     12.309      0.000       1.651       2.285
==============================================================================
Omnibus:                       34.113   Durbin-Watson:                   1.402
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               85.652
Skew:                           1.074   Prob(JB):                     2.52e-19
Kurtosis:                       6.538   Cond. No.                     6.04e+07
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 6.04e+07. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Exhibit C

```
                            OLS Regression Results
==============================================================================
Dep. Variable:         log_Concession   R-squared:                       0.844
Model:                            OLS   Adj. R-squared:                  0.840
Method:                 Least Squares   F-statistic:                     209.9
Date:                Fri, 11 Oct 2024   Prob (F-statistic):           1.07e-46
Time:                        22:48:29   Log-Likelihood:                 74.616
No. Observations:                 120   AIC:                            -141.2
Df Residuals:                     116   BIC:                            -130.1
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         14.2678      0.051    277.849      0.000      14.166      14.370
Parking     4.226e-08   4.97e-09      8.502      0.000    3.24e-08    5.21e-08
Rental Car  5.578e-08   1.16e-08      4.820      0.000    3.29e-08    7.87e-08
Ground      3.028e-07   4.19e-08      7.220      0.000     2.2e-07    3.86e-07
==============================================================================
Omnibus:                       39.377   Durbin-Watson:                   1.275
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              172.781
Skew:                          -1.011   Prob(JB):                     3.03e-38
Kurtosis:                       8.520   Cond. No.                     6.04e+07
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 6.04e+07. This might indicate that there are
strong multicollinearity or other numerical problems.
```

# Exhibit D

```
      feature        VIF
0       const  18.118450
1     Parking   1.818270
2  Rental Car   2.003398
3      Ground   2.124676
                            OLS Regression Results
==============================================================================
Dep. Variable:             Concession   R-squared:                       0.705
Model:                            OLS   Adj. R-squared:                  0.700
Method:                 Least Squares   F-statistic:                     140.0
Date:                Fri, 11 Oct 2024   Prob (F-statistic):           9.24e-32
Time:                        22:48:31   Log-Likelihood:                -1794.0
No. Observations:                 120   AIC:                             3594.
Df Residuals:                     117   BIC:                             3602.
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const        3.05e+05    2.9e+05      1.051      0.295    -2.7e+05     8.8e+05
Parking        0.1959      0.026      7.397      0.000       0.143       0.248
Rental Car     0.4478      0.059      7.621      0.000       0.331       0.564
==============================================================================
Omnibus:                       17.132   Durbin-Watson:                   0.891
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               24.108
Skew:                           0.731   Prob(JB):                     5.82e-06
Kurtosis:                       4.638   Cond. No.                     5.91e+07
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 5.91e+07. This might indicate that there are
strong multicollinearity or other numerical problems.
```

To improve R square and adjusted R square value Feature Engineering: Adding interaction terms or polynomial terms. Regularization: Using Ridge or Lasso to stabilize and regularize the model. Remove Insignificant Predictors: Drop predictors with high p-values. Outlier Detection: Remove influential points using Cook's Distance. Log Transformation: Apply log transformation to reduce skewness in the target variable. Add More Features: Incorporate additional relevant features to improve the explanatory power of the model.