**NYC Flights 2013: Exploratory Data Analysis Practice**

**Dataset Description**

A dataset of 336,776 domestic commercial flights from NYC's three major airports (Newark, JFK, and LaGuardia) in 2013. Available in CSV format or via the nycflights13 R package.

**Practice Activities**

1. **Goal Setting**

Define a SMART goal for your analysis.

Example: "Identify the top three factors contributing to flight delays at NYC airports in 2013 and propose data-driven solutions to reduce average delay times by December 15th."

2. **Variable Selection and Justification**

Choose variables that directly support your analysis goal.

Example:

* Primary outcome: Departure delay (minutes)
* Key predictors:
  + Carrier (airline performance)
  + Time of day (hourly patterns)
  + Weather conditions (environmental impact)

3. **Define Success Metrics**

Establish concrete measures for evaluating your analysis.

Example:

* Explain ≥60% of delay variance
* Develop 3+ actionable recommendations
* Create a prediction model with R² >0.7

4. **Business Context**

Connect your analysis to organizational goals.

Example:

* Improve customer satisfaction
* Reduce operational costs
* Optimize resource allocation

5. **Identify Constraints**

List key limitations.

Example:

* Analysis deadline: 2 weeks
* Missing data: 5% of weather records
* Computing power limitations

6. **Data Visualization Strategy**

Plan how to communicate findings effectively.

Example:

* Delay patterns: Heat maps by hour/month
* Airline comparison: Box plots
* Key factors: Decision trees

7. **Analysis Methods**

**Univariate Analysis**

Examine individual variables using Python.

Example:

A screen shot of a computer program

AI-generated content may be incorrect.

**Bivariate Analysis**

Study relationships between variables.

Example:

A screenshot of a computer program

AI-generated content may be incorrect.

8. **Dimensionality Reduction**

Apply and interpret PCA using scikit-learn.

Example:

A screen shot of a computer program

AI-generated content may be incorrect.

**Deliverable Format**

For each section, provide:

1. Analysis code and output
2. Clear visualizations
3. Written insights
4. Practical recommendations

**Sample Write-up Structure**

**A screenshot of a computer program

AI-generated content may be incorrect.**