

Dataset 1: Airlines Delays (airlines-delays.xlsx) This dataset includes comprehensive data on airline delays by cause—weather, carrier issues, security, and late aircraft—categorized by airline, airport, and month. It also details the number of flights, cancellations, and delays. Here are some suggested analyses using Advanced Analysis with Python in Copilot:

1. Forecast future airline delays at Albany International using the 'arr_delay' data.
2. Conduct a textual analysis of airport names to identify regional patterns in delays.
3. Generate a pairplot to visualize relationships between 'arr_delay', 'carrier_ct', and 'weather_ct'.
4. Create a jointplot to compare 'arr_delay' with 'weather_delay' using Seaborn.

Dataset 2: Yelp Reviews (yelp-reviews.xlsx) This dataset contains Yelp reviews with detailed attributes including business ID, review date, star ratings, and engagement metrics such as "cool," "useful," and "funny" votes. Suggested analyses include:

1. Analyze the sentiment of review texts and visualize the relationship between sentiment and star ratings.
2. Perform topic modeling using Latent Dirichlet Allocation (LDA) to identify prevalent topics within the reviews.
3. Visualize a heatmap to explore correlations between star ratings and engagement metrics.
4. Compare the length of reviews to the number of 'useful' votes using a jointplot.

Dataset 3: Employee Attrition (employee-attrition.xlsx) Focusing on employee attrition, this dataset includes demographic, job-related, and performance metrics, along with attrition status. Analytical suggestions include:

1. Develop a logistic regression model to predict attrition based on job-related features.
2. Employ a random forest model to determine feature importance for attrition prediction.
3. Display a heatmap of correlations among features such as years at the company, work-life balance, and attrition rates.
4. Create a stripplot to illustrate the distribution of work-life balance across different attrition statuses.

These datasets and analytical tasks are designed to not only practice your skills but also to deepen your understanding of how Python and Copilot can be leveraged for advanced analytical tasks beyond traditional Excel capabilities.

