QMB 6315: Python for Business Analytics
College of Business
University of Central Florida
Spring 2025

Assignment 4

Due Sunday, April 20, 2025 at 11:59 PM
in digital form in your GitHub repository.

Instructions:

Complete this assignment within the space on your *private* GitHub repo (not a fork of the course repo QMB6315S25) in a folder called `assignment_04`. In this folder, save your answers to Questions 1 to 3 in a file called `airplane_SQL.py`, following the sample script in the folder `assignment_04` in the course repository. When you are finished, submit it by uploading your files to your GitHub repo using the approach outlined in Question 4. You are free to discuss your approach to each question with your classmates but you must upload your own work.

Question 1:

The folder `assignment_04` contains a database `airplanes.db` with three tables: `Sales`, `Specs`, and `Perf`. The first table `Sales` contains the following variables.

`sale_id`	=	a unique key for each airplane sold
`age`	=	age of the aircraft, in years
`price`	=	price of an airplane

Use the first table, `Sales`, to estimate a regression model to predict the prices of airplanes.

a) Wrte a query that will obtain the data from the `Sales` table.

b) Obtain the data from the `Sales` table and store it in a data frame called `airplane_sales` in your workspace.

c) Estimate a regression model to predict `price` as a function of `age`. Inspect the printed estimation output from the command `print(reg_model_sales.summary())` as a second form of verification.

Question 2:

The next table `Specs` contains the following variables.

`sale_id`	=	a unique key for each airplane sold
`pass`	=	the number of passengers an airplane can accommodate
`wtop`	=	an indicator that the wings are above the fuselage
`fixgear`	=	an indicator for fixed landing gear (i.e. wheels are not retractable)
`tdrag`	=	an indicator that a wheel is on the tail (a tail-dragger)

Use both tables, `Sales` and `Specs`, to estimate a better regression model to predict the prices of airplanes.

a) Wrte a query that will obtain the data from the `Sales` table and join it to the data from the `Specs` table.

b) Obtain the data from the above query and store it in a data frame called `airplane_sales_specs` in your workspace.

c) Estimate a regression model to predict `price` as a function of `age`, `passengers`, `wtop`, `fixgear`, and `tdrag`. Inspect the printed estimation output from the command `print(reg_model_sales_specs.summary())` as a second form of verification.

Question 3:

The next table `Perf` contains the following variables.

`sale_id`	=	a unique key for each airplane sold
`horse`	=	the horsepower of the engine
`fuel`	=	the volume of the fuel tank, in gallons
`ceiling`	=	the maximum flying height of an airplane, in feet
`cruise`	=	the cruising speed, in MPH

Use all three tables to estimate an even better regression model to predict the prices of airplanes.

a) Wrte a query that will obtain the data from the `Sales` table and join it to the data from the `Specs` table and the the `Perf` table.

b) Obtain the data from the above query and store it in a data frame called `airplane_full` in your workspace.

c) Estimate a regression model to predict `price` as a function of `age`, `passengers`, `wtop`, `fixgear`, and `tdrag`, as well as `horse`, `fuel`, `ceiling`, and `cruise`. Inspect the printed estimation output from the command `print(reg_model_full.summary())` as a second form of verification.

Question 4:

Upload your code to your GitHub repo using the interface in GitHub Desktop.

1. Save your file within the folder in your repository in GitHub Desktop.
2. When you see the changes in GitHub Desktop, add a description of the changes you are making in the bottom left panel.
3. Press the button "Commit to main" to commit those changes.
4. Press the button "Push origin" to push the changes to the online repository. After this step, the changes should be visible on a browser, after refreshing the page.