Case Study: Self-Guided Version

In this version of the case study, you receive the high-level requirements to solve the business problem. There are multiple solutions to the problem, and how you solve it is your decision.

To solve the business problem, follow the requirements below given to you by your supervisor. Be aware that these requirements are only assumed for this case study. They are not an accurate representation of TSA or FAA requirements.

Begin the case study by opening the **StarterProgram.sas** and accessing the tables.

Deliverables and Requirements

Your job is to prepare two tables for analysis. After the tables are prepared, you can run the provided code in **AnalysisProgram.sas** to analyze the results. For the analysis program to run correctly, follow the requirements for each deliverable listed below.

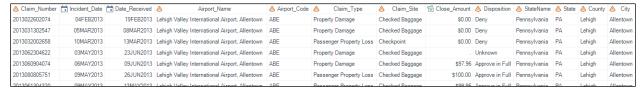
To create the following deliverables, you can use a variety of different methods that can include temporary or permanent tables, views, and in-line views. Be sure to explore the tables and columns and compare with the "Data Layout" section for all column requirements.

- 1. **sq.Claims_Cleaned** Create a new table named **sq.Claims_Cleaned** that cleans and prepares the **sq.claimsraw** table. Here is what the **sq.Claims_Cleaned** table must do:
 - a. Include all columns from the **sq.claimsraw** table and remove duplicated rows.
 - b. Change all missing values to *Unknown* for the following columns: **Airport_Code**, **Claim_Type**, **Claim_Site**, and **Disposition**. Follow the requirements in the "Data Layout" section for the column values.
 - c. Fix all rows where **Incident_Date** occurs **after Date_Received** by adding one year to the **Date_Received** value.
 - d. **StateName**, **City**, and **County** values should be in proper case (for example, *Raleigh*).
 - e. State values should be in uppercase.
 - f. Include only those rows where **Incident_Date** is between 2013 and 2017.
 - g. Currency columns should be permanently formatted with a dollar sign and include two decimal places (for example, \$130.28).
 - h. All dates should be permanently formatted in the style 01JAN2000.
 - Assign permanent labels for columns by adding a space between words (for example, Close Amount).
 - j. Order the final table by **Airport_Code** and **Incident_Date**.

Log

NOTE: Table TSA.CLAIMS CLEANED created, with 42522 rows and 13 columns.

Partial Table

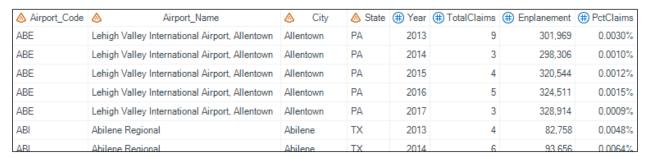


- sq.ClaimsByAirport Create a new table named sq.ClaimsByAirport by summarizing claims
 for each airport and year from the sq.Claims_Cleaned table. Then perform an inner join on the
 summarized data with the sq.enplanement2017 and sq.boarding2013_2016 tables.
 - a. Include the following columns from sq.Claims_Cleaned table: Airport_Code,
 Airport_Name, City, State, and the year of the Incident_Date. Name the new column Year.
 - b. Three new columns need to be added.
 - 1) Create the column **TotalClaims** as the number of claims for each group.
 - Retrieve the total passengers boarding for each Year and Airport_Code and name the column Enplanement. The information can be found in the sq.enplanement2017 and sq.boarding2013_2016 tables.
 - 3) Calculate the percentage of claims for each group by dividing **TotalClaims** by **Enplanement**. Name the new column **PctClaims** and format accordingly.
 - 4) Order the results by Airport_Code and Year.

Log

NOTE: Table TSA.CLAIMSBYAIRPORT created, with 1438 rows and 8 columns.

Partial Table



After you have prepared the data deliverables, open and run the AnalysisProgram.sas code to create
 FinalReport.html. Note: You must have the final tables in the Sq library for the program to run
 correctly. FinalReport.html is created in your course code folder. Use this report to answer the
 quiz questions.