CASE STUDY 1: ANALYZE UNITED STATES TSA CLAIMS DATA

CASE STUDY OVERVIEW

BUSSINESS SCENARIO

You've been asked to prepare and analyze data from the Transportation Security Administration (TSA). The TSA is an agency of the US Department of Homeland Security that has authority over the security of the traveling public.

Claims are filed if <u>travelers are injured, or their property is lost or damaged</u> during the screening process at an airport. You have data claims files <u>between 2002 and 2017</u>.

DATA INFORMATION

The case study data is in a **CSV file** (*TSAClaims2002_2017*), created from publicly available data from the TSA and the Federal Aviation Administration. The TSA data has information about the claims, and the FAA data has information about USA airport facilities.

The TSA data was created by <u>concatenating</u> TSA Claims Tables, removing certain columns, and joining the TSA and FAA data.

The CSV file has <u>14 columns</u> and over <u>220,855 rows</u>.

Columns:

- Claim_Number; some claims have duplicate claim numbers, but different info.
- Incident_Date, Date_Received; date of the incident, and date when the claim was filed.
- Claim_Type; 14 valid claim types.
- Claim_Site; 8 valid claim sites.
- Disposition; final settlement of the claim.
- Close_Amount; dollar amount of the settlement.
- Item_Category; type of items in the claim. The values in this column vary by year (we won't use this column).
- Airport_Code, Airport_Name;
- County, City, State, Statename; the State column has a 2 letter state code, and the State column contains the full name.