

```
libname tsa "/home/u63403183";
options validvarname=v7;

proc import datafile="/home/u63403183/EPG1V2/data/TSAClaims2002_2017.csv"
    out=tsa.ClaimsImport dbms=csv replace;
    guessingrows=max;
run;

/* Preview the data. */
proc print data=tsa.ClaimsImport (obs=20);
run;

proc contents data=tsa.ClaimsImport varnum;
run;

/* Change missing values and "-" to Unknown. */
/* Change Passenger Property Loss/Injury values to Passenger Property Loss. */
/* Change Property Damage/Personal Injury values to Property Damage. */
/* Remove the leading space in Closed: Canceled.. */
/* Fix the missing character C and the leading space in losed: Contractor Claim. */
/* Notice that many dates are after the year 2017 or missing. */
/* Notice that many dates are before 2002, after 2017, missing, or after Date_Received. */
proc freq data=tsa.Claims_NoDups;
    tables Claim_Site Disposition Claim_Type / nocum nopercnt;
    tables Date_Received Incident_Date / nocum nopercnt;
    format Date_Received Incident_Date year4.;
run;

/* Remove duplicate rows. */
proc sort data=tsa.ClaimsImport
out=tsa.Claims_NoDups
nodupkey;
by _all_;
run;

/* Sort the data by ascending Incident_Date. */
proc sort data=tsa.Claims_NoDups;
    by Incident_Date;
run;

data tsa.Claims_NoDups;
    set tsa.Claims_NoDups;

    /* Clean the Claim_Site column. */
    if Claim_Site in ('-', '') then
        Claim_Site="Unknown";

    /* Clean the Disposition column. */
    if Disposition in ("-", "") then
        Disposition='Unknown';
    else if Disposition='Closed: Canceled' then
        Disposition='Closed:Canceled';
    else if Disposition='losed: Contractor Claim' then
        Disposition='Closed:Contractor Claim';

    /* Clean the Claim_Type column. */
    if Claim_Type in ("-", "") then
        Claim_Type="Unknown";
    else if Claim_Type='Passenger Property Loss/Personal Injur' then
        Claim_Type='Passenger Property Loss';
    else if Claim_Type='Passenger Property Loss/Personal Injury' then
```

```

        Claim_Type='Passenger Property Loss';
    else if Claim_Type='Property Damage/Personal Injury' then
        Claim_Type='Property Damage';
run;

data tsa.Claims_NoDups;
    set tsa.Claims_NoDups;

    /* Convert all State values to uppercase and all StateName values to proper case. */
    State=upcase(state);
    StateName=propcase(StateName);
run;

data tsa.Claims_NoDups;
    set tsa.Claims_NoDups;
/* Create a new column that indicates date issues. */
if (Incident_Date > Date_Received or Incident_Date=. or Date_Received=. or
    year(Incident_Date) < 2002 or year(Incident_Date) > 2017 or
    year(Date_Received) < 2002 or year(Date_Received) > 2017) then
    Date_Issues="Needs Review";
run;

data tsa.Claims_NoDups;
    set tsa.Claims_NoDups;
/* Add permanent labels and formats. */
format Incident_Date Date_Received date9. Close_Amount Dollar20.2;
label Airport_Code="Airport Code" Airport_Name="Airport Name"
    Claim_Number="Claim Number" Claim_Site="Claim Site" Claim_Type="Claim Type"
    Close_Amount="Close Amount" Date_Issues="Date Issues"
    Date_Received="Date Received" Incident_Date="Incident Date"
    Item_Category="Item Category";

/* Drop County and City. */
drop County City;
run;

%let outpath= /home/u63403183;
ODS PDF FILE="&outpath/ClaimsReports.pdf" STYLE=Meadow STARTPAGE=NO PDFTOC=1;
ODS PROCLABEL"TSAreport";
title "Overall Date Issues in the Data";

/* Analyze the overall data to answer the business questions. Be sure to add appropriate titles. */
title "Overall Date Issues in the Data";
proc freq data=tsa.Claims_NoDups;
table Date_Issues / nocum nopercent;
run;
title;
ods graphics on;
title "Overall Claims by Year";

proc freq data=tsa.Claims_NoDups;
table Incident_Date / nocum nopercent plots=freqplot;
format Incident_Date year4.;
where Date_Issues is null;
run;
title;

%let StateName=California;
title "&StateName Claim Types, Claim Sites and Disposition
Frequencies";
proc freq data=tsa.Claims_NoDups order=freq;

```

```
table Claim_Type Claim_Site Disposition / nocum nopercnt;  
where StateName="&StateName" and Date_Issues is null;  
run;  
title "Close_Amount Statistics for &StateName";  
proc means data=tsa.Claims_NoDups mean min max sum maxdec=0;  
var Close_Amount;  
where StateName="&StateName" and Date_Issues is null;  
run;  
title;  
ODS PDF CLOSE;
```