```
%let path = /home/u64076587/ECRB94/data;
libname tsa "&path";

options validvarname=v7;

proc import datafile="&path/TSAClaims2002_2017.csv"
            dbms=csv
            out=tsa.ClaimsImport
            replace;
        guessingrows=max;
run;
proc print data=tsa.ClaimImport (obs=20);
run;
proc contents data=tsa.ClaimsImport varnum;
run;
proc freq data=tsa.ClaimImport;
    tables claim_site
            disposition
            claim_type
            date_received
            incident_date / nocum nopercent;
        format incident_date date_received year4.;
run;

proc print data=tsa.ClaimsImport;
    where date_received < incident_date;
    format date_received incident_Date date9.;
run;

/*removing duplicates*/
proc sort data=tsa.ClaimsImport
            out=tsa.Claims_NoDups noduprecs;
    by _all_;
run;

/*SORT DATE BY INCIDENT DATE*/
proc sort data=tsa.Claims_NoDups;
    by incident_Date;
run;

data tsa.claims_cleaned;
    set tsa.claims_nodups;

/*Clean calim_site col*/
    if Claim_Site in('-','') then Claim_Site="Unknown";

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

/*clean calim_type col*/
    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';

/*convert all states values to uppercase and all state name to proper case*/
    State=upcase(state);
    StateName=propcase(StateName);
/*new col to indiacte date issues*/
    if (Incident_Date > Date_Received or
        Date_Received = . or
        Incident_Date = . 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";

/*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_Anount="Close Amount"
          Date_Issues="Date Issues"
```

```
          Date_Received="Date Received"
          Incident_Date="Incident Date"
          Item_Category="Item Category";

/*drop country and city*/
    drop county city;
run;

proc freq data=tsa.Claims_Cleaned order=freq;
    tables Claim_Site
        Disposition
        Claim_Type
        Date_Issues / nopercent nocum;
run;


%let statename=Hawaii;

title "Overall Date Issues in the data";
proc freq dat=TSA.Claims_Cleaned;
    table Date_Issues /nocum nopercent;
run;
title;

ods graphics on;
title "Overall Claims By Year";
proc freq data=TSA.Claims_Cleaned;
    table Incident_Date /nocum nopercent plots=freqplot;
    format Incident_Date year4.;
    where Date_Issues is null;
run;
title;
%let statename=Hawaii;
title "Hawaii Claim Types, Claim Sites and Disposition";
proc freq data=TSA.Claims_Cleaned order=freq;
    table Claim_Type Claim_Site Disposition / nocum nopercent;
    where StateName="&StateName" and Date_Issues is null;
run;
title;

title "Close Amount Sta for Hawaii";
proc means data=TSA.Claims_Cleaned mean min max sum maxdec=0;
    var Close_Amount;
    where StateName="&StateName" and Date_Issues is null;
run;
title;
```