

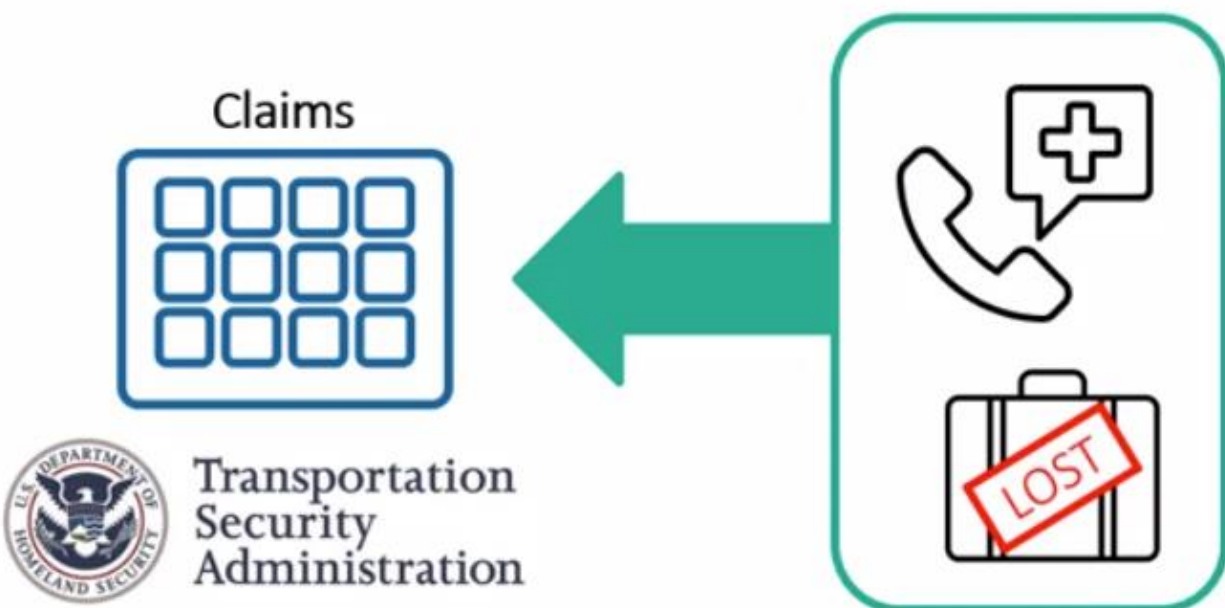
## Practice for the SAS Programmer Certification

### Case Study 1: TSA Claims

# Analyze United States TSA Claims Data

SAS Programming Case Study

## Business Scenario



## Data Information

TSAClaims2002\_2017



## TSAClaims2002\_2017



14 Columns

220,855 Rows

Claim\_Number

Incident\_Date, Date\_Received

Claim\_Type

Claim\_Site

Disposition

Close\_Amount

Item\_Category

Airport\_Code, Airport\_Name

County, City, State, Statename

Access  
data

Explore  
data

Prepare  
data

Analyze and  
report on  
data

Export  
results



PRINT

MEANS

FREQ



```
/* Accessing Data */
```

```
%let path=/home/u58304328/ECRB94/data;
```

```
libname tsa "&path";
```

```
options validvarname=v7;
```

```
proc import datafile="&path/TSAClaims2002_2017.csv"
```

```
    dbms=csv
```

```
    out=tsa.ClaimsImport
```

```
    replace;
```

```
    guessingrows=max;
```

```
run;
```

Table: TSA.CLAIMSIMPORT View: Column names Filter: (none)

Columns

	Claim_Number	Date_Received	Incident_Date	Airport_Code	Airport_Name	Claim_Type	Claim_Site
1	2006081611123	17027	17006			Passenger Property Loss	Checked B
2	2006062108380	16972	16957			Passenger Property Loss	Checkpoint
3	2006062008258	16972	16938			Passenger Property Loss	Checked B
4	2006010699056	16831	16793			Passenger Property Loss	Checked B
5	2006032303625	16880	16861			Property Damage	Checked B
6	2006081511005	17021	.			Property Damage	Checked B
7	2006082111318	17029	17015			Property Damage	Checked B
8	2006070709101	16988	16979			Passenger Property Loss	Checkpoint
9	2006072109924	17003	16996			Passenger Property Loss	Checkpoint
10	2006041704637	16898	.			Passenger Property Loss	Checked B
11	2006081010689	17016	16866			Property Damage	Checked B
12	2006042505214	16905	16896			Property Damage	Checkpoint
13	2006061408002	16966	.				
14	2006081110819	17022	17014			Passenger Property Loss	Checked B

```
/* Exploring Data */
```

```
proc print data=tsa.ClaimsImport(obs=30);
```

```
run;
```

Obs	Claim_Number	Date_Received	Incident_Date	Airport_Code	Airport_Name	Claim_Type	Claim_Site	Item_Category	Close_Amount	Disposition	StateName	State	County	City
1	2008081811123	17027	17006			Passenger Property Loss	Checked Baggage	Candles - Decorative and other; Clothing - Shoes, belts, accessories, etc.; Dishes, Pottery, Glassware, Plasticware	-					
2	2008082108380	16972	16957			Passenger Property Loss	Checkpoint	Jewelry - Fine	-					
3	2008082008258	16972	16938			Passenger Property Loss	Checked Baggage	Cosmetics - Perfume, toilet articles, medicines, soaps, etc.; Medicines	-					
4	2008010690056	16831	16793			Passenger Property Loss	Checked Baggage	Other	-					
5	2008032303625	16880	16861			Property Damage	Checked Baggage	Luggage (all types including footlockers)	-					
6	2008081511005	17021	-			Property Damage	Checked Baggage	Looks: Luggage (all types including footlockers)	-					
7	2008082111318	17029	17015			Property Damage	Checked Baggage	Looks	-					
8	2008070709101	16968	16979			Passenger Property Loss	Checkpoint	Jewelry - Fine	-					
9	2008072109924	17003	16996			Passenger Property Loss	Checkpoint	DVD/CD Players: Other	-					
10	2008041704837	16968	-			Passenger Property Loss	Checked Baggage	Eyeglasses - (including contact lenses)	-					
11	2008081010689	17016	16896			Property Damage	Checked Baggage	Luggage (all types including footlockers)	-					
12	2008042505214	16905	16896			Property Damage	Checkpoint	Cell Phones	-					
13	2008081408002	16966	-					Luggage (all types including footlockers)	-					
14	2008081110819	17022	17014			Passenger Property Loss	Checked Baggage	Other	-					
15	2008051606567	16937	16924			Property Damage	Checked Baggage	Electrical and Gas Appliances Minor - \$200 or less (humidifiers, tv's, etc)	-					
16	2008040504066	16877	16855			Passenger Property Loss	Checked Baggage	Other	-					
17	2008073110251	17008	17001			Passenger Property Loss	Checked Baggage	Clothing - Shoes, belts, accessories, etc.	-					
18	2008042505180	16909	16883			Passenger Property Loss	Checkpoint	Eyeglasses - (including contact lenses)	-					
19	2008020801205	16834	16828			Passenger Property Loss	Checked Baggage	Cameras - Digital	0	Deny				
20	2008030302477	16859	16829			Passenger Property Loss	Checkpoint		-					
21	2008032303589	16880	-			Passenger Property Loss	Checked Baggage	Jewelry - Fine	-					

```
proc contents data=tsa.ClaimsImport varnum;
```

```
run;
```

### Observation results:

1. There are 220,855 records with 14 variables (columns)
2. Date\_Received and Incident\_Date formats are BEST12, where it should be DATE format
3. Close\_Amount format is BEST12, where it should be dollar amount format

### The CONTENTS Procedure

Data Set Name	TSA.CLAIMSIMPORT	Observations	220855
Member Type	DATA	Variables	14
Engine	V9	Indexes	0
Created	05/09/2021 21:21:28	Observation Length	1072
Last Modified	05/09/2021 21:21:28	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	1811
First Data Page	1
Max Obs per Page	122
Obs in First Data Page	118
Number of Data Set Repairs	0
Filename	/home/u58304328/ECRB94/data/claimsimport.sas7bdat
Release Created	9.0401M6
Host Created	Linux
Inode Number	12962729974
Access Permission	rw-r--r--
Owner Name	u58304328
File Size	227MB
File Size (bytes)	237502464

Variables in Creation Order					
#	Variable	Type	Len	Format	Informat
1	Claim_Number	Char	13	\$13.	\$13.
2	Date_Received	Num	8	BEST12.	BEST32.
3	Incident_Date	Num	8	BEST12.	BEST32.
4	Airport_Code	Char	3	\$3.	\$3.
5	Airport_Name	Char	48	\$48.	\$48.
6	Claim_Type	Char	39	\$39.	\$39.
7	Claim_Site	Char	15	\$15.	\$15.
8	Item_Category	Char	834	\$834.	\$834.
9	Close_Amount	Num	8	BEST12.	BEST32.
10	Disposition	Char	23	\$23.	\$23.
11	StateName	Char	17	\$17.	\$17.
12	State	Char	2	\$2.	\$2.
13	County	Char	20	\$20.	\$20.
14	City	Char	33	\$33.	\$33.

```

proc freq data=tsa.ClaimsImport;

    tables claim_site

        disposition

        claim_type

        date_received

        incident_date / nocum nopercnt;

    format date_received incident_date year4.;

run;

```

### Observation Results:

1. Claim\_Site has 387 missing values (-) which should be replaced with UNKNOWN values
2. Disposition has 12,766 missing values (-) which should be replaced with UNKNOWN values
3. Disposition has multiple variation of "Closed: Canceled" values which should be "Closed: Canceled"
4. Disposition has misspelled "losed: Contractor Claim: values which should be "Closed: Contractor Claim" values
5. Claim\_Type has 375 missing values (-) which should be replaced with UNKNOWN values
6. Claim\_Type has multiple variation of "Passenger Property Loss" values which should be "Passenger Property Loss"
7. Claim\_Type has multiple variation of "Property Damage" values which should be "Property Damage"
8. Date Received has 260 missing values
9. Date\_Received has invalid dates that are beyond year 2017
10. Incident\_Date has 2,329 missing values
11. Incident\_Date has invalid dates that are beyond year 2002 to 2017

### The FREQ Procedure

Claim_Site	Frequency
-	387
Bus Station	20
Checked Baggage	171504
Checkpoint	44748
Motor Vehicle	540
Not Provided	1
Other	2912
Pre-Check	8
Frequency Missing = 735	

Disposition	Frequency
*Insufficient	1735
-	12768
Approve in Full	48956
Closed: Canceled	167
Closed:Canceled	284
Closed:Contractor Claim	115
Deny	99600
In Review	8938
Pending Payment	1
Received	14
Settle	34422
losed: Contractor Claim	73
Frequency Missing = 13784	

Claim_Type	Frequency
-	375
Bus Terminal	1
Complaint	73
Compliment	3
Employee Loss (MPCECA)	492
Missed Flight	32
Motor Vehicle	423
Not Provided	2
Passenger Property Loss	126795
Passenger Property Loss/Personal Injur	8
Passenger Property Loss/Personal Injury	13
Passenger Theft	479
Personal Injury	1613
Property Damage	82603
Property Damage/Personal Injury	14
Property Loss	17
Wrongful Death	4
Frequency Missing = 7908	

Date_Received	Frequency
2002	1054
2003	22474
2004	29778
2005	24539
2006	19091
2007	18670
2008	16439
2009	12803
2010	10925
2011	10905
2012	10069
2013	9698
2014	8855
2015	8667
2016	7973
2017	8618
2020	1
2025	5
2040	1
2044	1
2055	1
2094	1
2200	1
2204	7
2205	4
2206	2
2207	1
2208	1
2500	2
2900	1
2907	1
2996	1

3004	1
3005	5
Frequency Missing = 260	

Incident_Date	Frequency
1996	1
2000	14
2001	15
2002	2158
2003	24508
2004	28960
2005	22971
2006	18806
2007	18275
2008	15902
2009	11179
2010	12215
2011	10995
2012	9997
2013	9536
2014	8680
2015	7721
2016	8188
2017	8403
2018	2
Frequency Missing = 2329	

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

### Observation Results:

There are data with Incident\_Date after Date\_Received that means date errors that are needed to be addressed.



Obs	Claim_Number	Date_Received	Incident_Date	Airport_Code	Airport_Name	Claim_Type	Claim_Site	Item_Category	Close_Amount	Disposition	StateName	State	County	City
74	2009021401524	08FEB2006	26DEC2006			Property Damage	Checkpoint	Computer - Laptop	-					
94	2009092706606	27JUN2006	10JUL2006			Passenger Property Loss	Checkpoint	Cameras - Digital	-					
284	2005091592532	08AUG2005	17AUG2005			Property Damage	Checkpoint	Other	-					
305	2005081090324	02AUG2005	07OCT2005			Property Damage	Checked Baggage	Luggage (all types including footlockers)	-					
328	2005121597898	08DEC2005	13DEC2005			Property Damage	Checked Baggage	Cosmetics - Perfume, toilet articles, medicines, soaps, etc.	138.74	Settle				
400	2005121297502	02DEC2005	18NOV2006			Passenger Property Loss	Checked Baggage		-					
409	2009011296526	04JAN2006	05DEC2006			Passenger Property Loss	Checked Baggage	Luggage (all types including footlockers)	-					
562	2005040781839	05APR2005	28SEP2005			Property Damage	Checked Baggage	Clothing - Shoes, belts, accessories, etc.	0	Deny				
582	2005052485505	23MAY2005	27DEC2005			Property Damage	Checked Baggage	Luggage (all types including footlockers)	-					
738	2005051985062	07MAR2005	11NOV2005			Passenger Property Loss	Checked Baggage	Other	-					
781	2005033180749	23MAR2005	21DEC2005			Property Damage	Checkpoint	Clothing - Shoes, belts, accessories, etc.	-					
804	2005042883426	20APR2005	25APR2005						0					
827	2004122971568	29OCT2004	25NOV2004			Passenger Property Loss	Checked Baggage	Jewelry - Fine	-					
1063	2009012600380	20JAN2005	09DEC2005			Property Damage	Checked Baggage	Luggage (all types including footlockers)	-					
1123	2004092763945	23SEP2004	10NOV2004			Property Damage	Checked Baggage	Other	-					
1224	2004122170045	01AUG2004	28SEP2004						0					
1269	2004070658310	17JUN2004	21JUN2004			Passenger Property Loss	Checked Baggage	Computer - Laptop	164.88	Approve in Full				
1303	2004102265617	09AUG2004	01SEP2004			Property Damage	Checked Baggage	Luggage (all types including footlockers)	50.09	Approve in Full				

/\* Preparing Data \*/

/\* 1. Remove Duplicate Rows \*/

```
proc sort data=tsa.ClaimsImport
```

```
    out=tsa.claims_cleaned nodupkey;
```

```
    by _all_;
```

```
run;
```

```
73      proc sort data=tsa.ClaimsImport
74          out=tsa.claims_NoDups nodupkey;
75          by _all_;
76      run;
```

NOTE: There were 220855 observations read from the data set TSA.CLAIMSIMPORT.

NOTE: 5 observations with duplicate key values were deleted.

NOTE: The data set TSA.CLAIMS\_NODUPS has 220850 observations and 14 variables.

```
proc sort data=tsa.claims_cleaned;
```

```
    by Incident_Date;
```

```
run;
```

```

78      /* 2. Sort the data by ascending Incident_Date */
79      proc sort data=tsa.claims_NoDups;
80      by Incident_Date;
81      run;

```

NOTE: There were 220850 observations read from the data set TSA.CLAIMS\_NODUPS.  
NOTE: The data set TSA.CLAIMS\_NODUPS has 220850 observations and 14 variables.

```

data tsa.claims_cleaned;

    set tsa.claims_NoDups;

/* 3. Clean the Claim_Site column */

    if Claim_Site in ('-', '') then Claim_Site="Unknown";

/* 4. Clean the Disposition column */

    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";

/* 5. 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";

/* 6. Convert All State values to Uppercase and all State_Name to proper case */

    State=upcase(state);

    StateName=propcase(StateName);

/* 7. Create a new column to indicate 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";
/* 8. Add permanent labels and format */
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";
/* 9. Drop County and City */
drop County City;
run;

NOTE: There were 220850 observations read from the data set TSA.CLAIMS_NODUPS.
NOTE: The data set TSA.CLAIMS_CLEANED has 220850 observations and 13 variables.

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

```

### The FREQ Procedure

Claim Site	
Claim_Site	Frequency
Checked Baggage	171503
Checkpoint	44748
Other	2912
Unknown	1122
Motor Vehicle	536
Bus Station	20
Pre-Check	8
Not Provided	1

Claim Type	
Claim_Type	Frequency
Passenger Property Loss	126816
Property Damage	82616
Unknown	8263
Personal Injury	1613
Employee Loss (MPCECA)	492
Passenger Theft	479
Motor Vehicle	419
Complaint	73
Missed Flight	32
Property Loss	17
Wrongful Death	4
Compliment	3
Not Provided	2
Bus Terminal	1

Disposition	Frequency
Deny	99800
Approve in Full	48956
Settle	34422
Unknown	26545
In Review	8938
*Insufficient	1735
Closed:Canceled	451
Closed:Contractor Claim	188
Received	14
Pending Payment	1

Date Issues	
Date_Issues	Frequency
Needs Review	4241
Frequency Missing = 216609	

/\* Overall Analysis \*/

/\* 1. How many date issues are in the overall data? \*/

title "Overall Data Issues in the Data";

proc freq data=tsa.claims\_cleaned;

table date\_issues / missing nocum nopercnt;

run;

title;

### Overall Data Issues in the Data

#### The FREQ Procedure

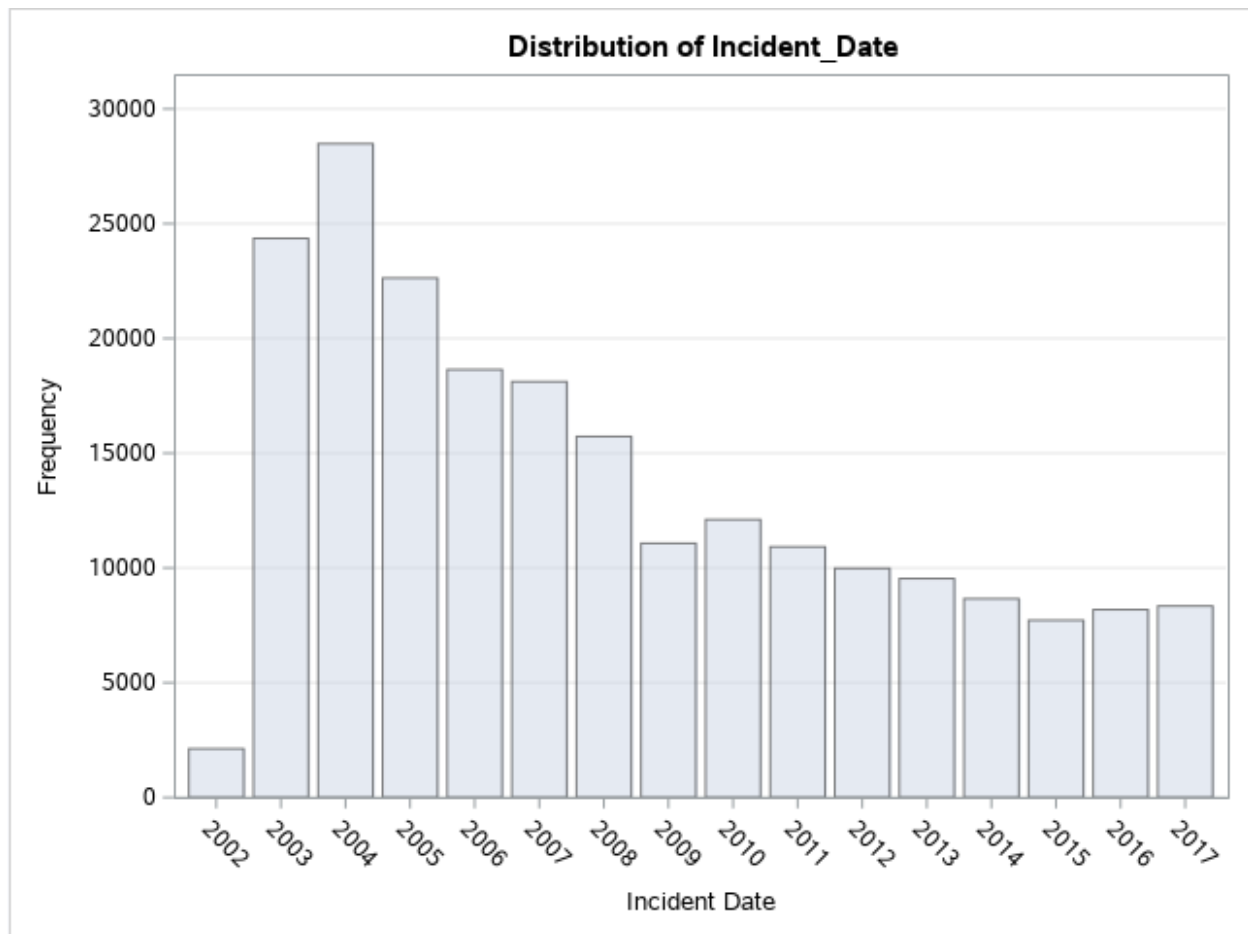
Date Issues	
Date_Issues	Frequency
	216609
Needs Review	4241

```
/* 2. How many claims per year of incident_date are in the overall data? Be sure to include a plot */  
ods graphics on;  
title "Overall Claims by Year";  
proc freq data=tsa.claims_cleaned;  
    table Incident_Date / nocum nopercents plots=freqplot;  
    format Incident_Date year4.;  
    where Date_Issues is null;  
run;  
title;
```

## Overall Claims by Year

The FREQ Procedure

Incident Date	
Incident_Date	Frequency
2002	2123
2003	24359
2004	28484
2005	22631
2006	18643
2007	18116
2008	15727
2009	11075
2010	12108
2011	10921
2012	9984
2013	9536
2014	8659
2015	7721
2016	8182
2017	8340



/\* Specific State Analysis \*/

/\* 3. Lastly, a user should be able to dynamically input a specific state value and answer the following questions \*/

/\* a. What are the frequency values for Claim\_Type for the selected state? \*/

/\* b. What are the frequency values for Claim\_Site for the selected state? \*/

/\* c. What are the frequency values for Disposition for the selected state? \*/

title "California Claim Type, Claim Site, and Disposition";

proc freq data=tsa.claims\_cleaned order=freq;

table Claim\_Type Claim\_Site Disposition / nocum noperc;

where StateName="California" and Date\_Issues is null;

run;

title;

## California Claim Type, Claim Site, and Disposition

### The FREQ Procedure

Claim Type	
Claim_Type	Frequency
Passenger Property Loss	14892
Property Damage	8996
Unknown	756
Personal Injury	194
Passenger Theft	55
Employee Loss (MPCECA)	51
Motor Vehicle	17
Complaint	7
Compliment	2
Missed Flight	1
Property Loss	1

Claim Site	
Claim_Site	Frequency
Checked Baggage	19553
Checkpoint	5142
Other	187
Unknown	68
Motor Vehicle	19
Pre-Check	2
Bus Station	1

Disposition	Frequency
Deny	10918
Approve in Full	5266
Settle	4192
Unknown	3188
In Review	1086
*Insufficient	187
Closed:Contractor Claim	83
Closed:Canceled	49
Received	3

/\* d. What is the mean, minimum, maximum, and sum for the close\_amount for the selected state?  
Round to the nearest Integer \*/

title "Close Amount Statistics for California";

proc means data=tsa.claims\_cleaned mean min max sum maxdec=0;

var Close\_Amount;



```

        where StateName="California" and Date_Issues is null;

run;

title;

```

Close Amount Statistics for California			
The MEANS Procedure			
Analysis Variable : Close_Amount Close Amount			
Mean	Minimum	Maximum	Sum
98	0	14519	2096386

```

%let statename=Texas;

```

```

/* Overall Analysis */

```

```

/* 1. How many date issues are in the overall data? */

```

```

title "Overall Data Issues in the Data";

```

```

proc freq data=tsa.claims_cleaned;

```

```

    table date_issues / missing nocum nopercnt;

```

```

run;

```

```

title;

```

```

/* 2. How many claims per year of incident_date are in the overall data? Be sure to include a plot */

```

```

ods graphics on;

```

```

title "Overall Claims by Year";

```

```

proc freq data=tsa.claims_cleaned;

```

```

    table Incident_Date / nocum nopercnt plots=freqplot;

```

```

    format Incident_Date year4.;

```

```

    where Date_Issues is null;

```

```

run;

```

```

title;

```

```
/* Specific State Analysis */
```

```
/* 3. Lastly, a user should be able to dynamically input a specific state value and answer the following questions */
```

```
/* a. What are the frequency values for Claim_Type for the selected state? */
```

```
/* b. What are the frequency values for Claim_Site for the selected state? */
```

```
/* c. What are the frequency values for Disposition for the selected state? */
```

```
title "&StateName Claim Type, Claim Site, and Disposition";
```

```
proc freq data=tsa.claims_cleaned order=freq;
```

```
    table Claim_Type Claim_Site Disposition / nocum nopercnt;
```

```
    where StateName="&StateName" and Date_Issues is null;
```

```
run;
```

```
title;
```

```
/* d. What is the mean, minimum, maximum, and sum for the close_amount for the selected state?  
Round to the nearest Integer */
```

```
title "Close Amount Statistics for &StateName";
```

```
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;
```

## Overall Data Issues in the Data

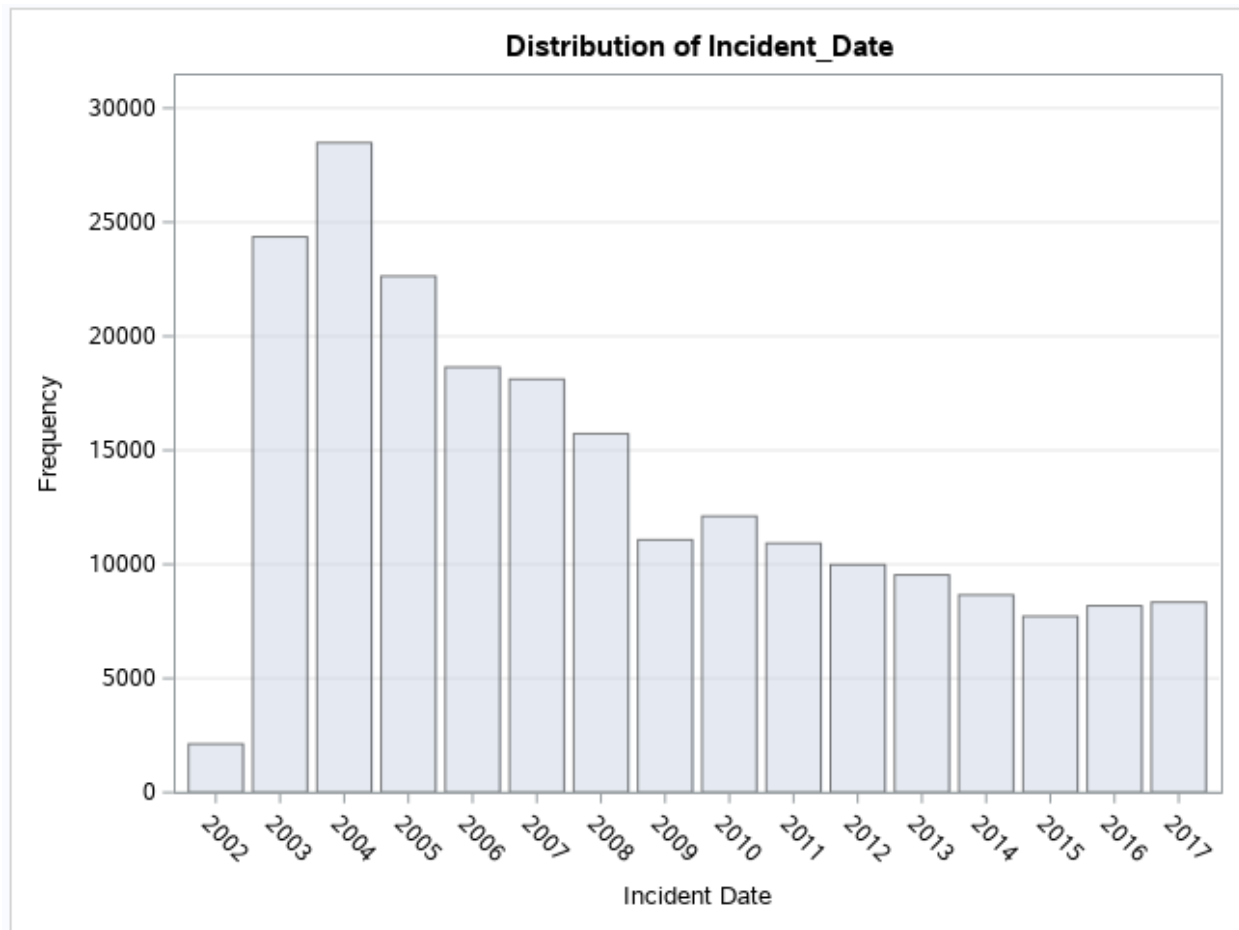
The FREQ Procedure

Date Issues	
Date_Issues	Frequency
	216609
Needs Review	4241

## Overall Claims by Year

The FREQ Procedure

Incident Date	
Incident_Date	Frequency
2002	2123
2003	24359
2004	28484
2005	22631
2006	18643
2007	18116
2008	15727
2009	11075
2010	12108
2011	10921
2012	9984
2013	9536
2014	8659
2015	7721
2016	8182
2017	8340



## Texas Claim Type, Claim Site, and Disposition

### The FREQ Procedure

Claim Type	
Claim_Type	Frequency
Passenger Property Loss	8093
Property Damage	5583
Unknown	343
Personal Injury	108
Employee Loss (MPCECA)	30
Passenger Theft	29
Motor Vehicle	8
Missed Flight	4
Complaint	2
Property Loss	2

Claim Site	
Claim_Site	Frequency
Checked Baggage	11139
Checkpoint	2945
Other	71
Unknown	38
Motor Vehicle	8
Pre-Check	1

Disposition	Frequency
Deny	6369
Approve in Full	3279
Settle	2297
Unknown	1334
In Review	753
*Insufficient	128
Closed:Canceled	37
Closed:Contractor Claim	4
Received	1

## Close Amount Statistics for Texas

### The MEANS Procedure

Analysis Variable : Close_Amount Close Amount			
Mean	Minimum	Maximum	Sum
101	0	106000	1225577

**/\* Exporting Result \*/**

%let statename=California;

%let outpath=/home/u58304328/ECRB94/output;

ods pdf file="&outpath/ClaimReports.pdf" style=meadow pdftoc=1;

ods noproctitle;

**/\* Overall Analysis \*/**

**/\* 1. How many date issues are in the overall data? \*/**

ods proclabel "Overall Date Issues";

title "Overall Data Issues in the Data";

proc freq data=tsa.claims\_cleaned;

    table date\_issues / missing nocum nopercnt;

run;

title;

**/\* 2. How many claims per year of incident\_date are in the overall data? Be sure to include a plot \*/**

ods graphics on;

ods proclabel "Overall Claims by Year";

title "Overall Claims by Year";

proc freq data=tsa.claims\_cleaned;

    table Incident\_Date / nocum nopercnt plots=freqplot;

    format Incident\_Date year4.;

    where Date\_Issues is null;

run;

title;

**/\* Specific State Analysis \*/**

/\* 3. Lastly, a user should be able to dynamically input a specific state value and answer the following questions \*/

/\* a. What are the frequency values for Claim\_Type for the selected state? \*/

/\* b. What are the frequency values for Claim\_Site for the selected state? \*/

/\* c. What are the frequency values for Disposition for the selected state? \*/

ods proclabel "&StateName Claims Overview";

title "&StateName Claim Type, Claim Site, and Disposition";

proc freq data=tsa.claims\_cleaned order=freq;

table Claim\_Type Claim\_Site Disposition / nocum nopercnt;

where StateName="&StateName" and Date\_Issues is null;

run;

title;

/\* d. What is the mean, minimum, maximum, and sum for the close\_amount for the selected state?  
Round to the nearest Integer \*/

ods proclabel "&StateName Close Amount Statistics";

title "Close Amount Statistics for &StateName";

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;

ods pdf close;

### Overall Data Issues in the Data

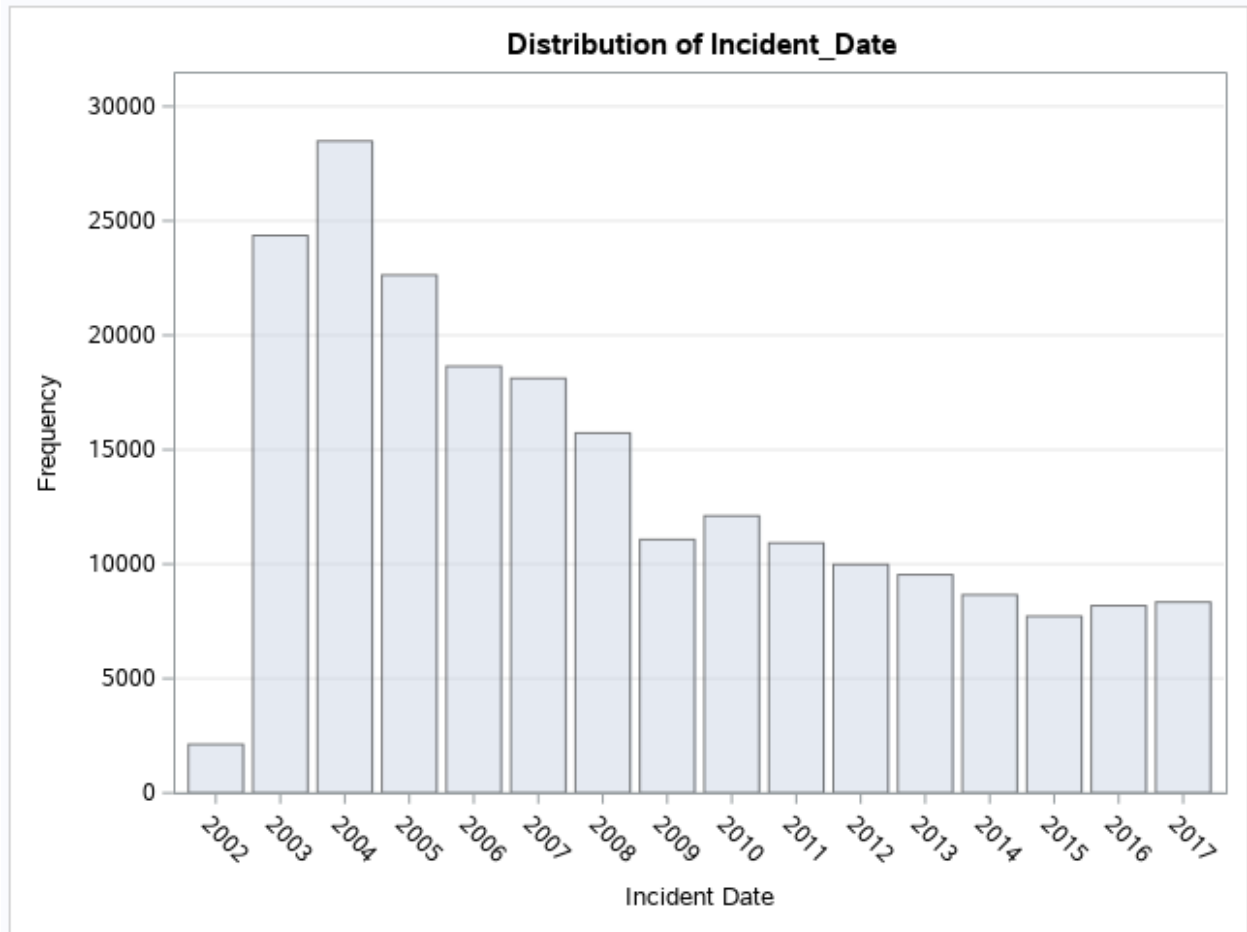
Date Issues	
Date_Issues	Frequency
	216609
Needs Review	4241

---

### Overall Claims by Year

Incident Date	
Incident_Date	Frequency
2002	2123
2003	24359
2004	28484
2005	22631
2006	18643
2007	18116
2008	15727
2009	11075
2010	12108
2011	10921
2012	9984
2013	9536
2014	8659
2015	7721
2016	8182
2017	8340





## California Claim Type, Claim Site, and Disposition

Claim Type	
Claim_Type	Frequency
Passenger Property Loss	14882
Property Damage	8996
Unknown	756
Personal Injury	194
Passenger Theft	55
Employee Loss (MPCECA)	51
Motor Vehicle	17
Complaint	7
Compliment	2
Missed Flight	1
Property Loss	1

Claim Site	
Claim_Site	Frequency
Checked Baggage	19553
Checkpoint	5142
Other	187
Unknown	68
Motor Vehicle	19
Pre-Check	2
Bus Station	1

Disposition	Frequency
Deny	10918
Approve in Full	5266
Settle	4192
Unknown	3188
In Review	1086
*Insufficient	187
Closed:Contractor Claim	83
Closed:Canceled	49
Received	3

## Close Amount Statistics for California

Analysis Variable : Close_Amount Close Amount			
Mean	Minimum	Maximum	Sum
98	0	14519	2096386

```

%let statename=Hawaii;

%let outpath=/home/u58304328/ECRB94/output;

ods pdf file="&outpath/ClaimReports.pdf" style=meadow pdftoc=1 ;

ods noproctitle;

/* Overall Analysis */

/* 1. How many date issues are in the overall data? */
ods proclabel "Overall Date Issues";
title "Overall Data Issues in the Data";
proc freq data=tsa.claims_cleaned;
    table date_issues / missing nocum nopercnt;
run;
title;

/* 2. How many claims per year of incident_date are in the overall data? Be sure to include a plot */
ods graphics on;
ods proclabel "Overall Claims by Year";
title "Overall Claims by Year";
proc freq data=tsa.claims_cleaned;
    table Incident_Date / nocum nopercnt plots=freqplot;
    format Incident_Date year4.;
    where Date_Issues is null;
run;
title;

/* Specific State Analysis */

/* 3. Lastly, a user should be able to dynamically input a specific state value and answer the following
questions */

```

```

/* a. What are the frequency values for Claim_Type for the selected state? */
/* b. What are the frequency values for Claim_Site for the selected state? */
/* c. What are the frequency values for Disposition for the selected state? */
ods proclabel "&StateName Claims Overview";
title "&StateName Claim Type, Claim Site, and Disposition";
proc freq data=tsa.claims_cleaned order=freq;
    table Claim_Type Claim_Site Disposition / nocum nopercnt;
    where StateName="&StateName" and Date_Issues is null;
run;
title;

/* d. What is the mean, minimum, maximum, and sum for the close_amount for the selected state?
Round to the nearest Integer */
ods proclabel "&StateName Close Amount Statistics";
title "Close Amount Statistics for &StateName";
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;

ods pdf close;

```

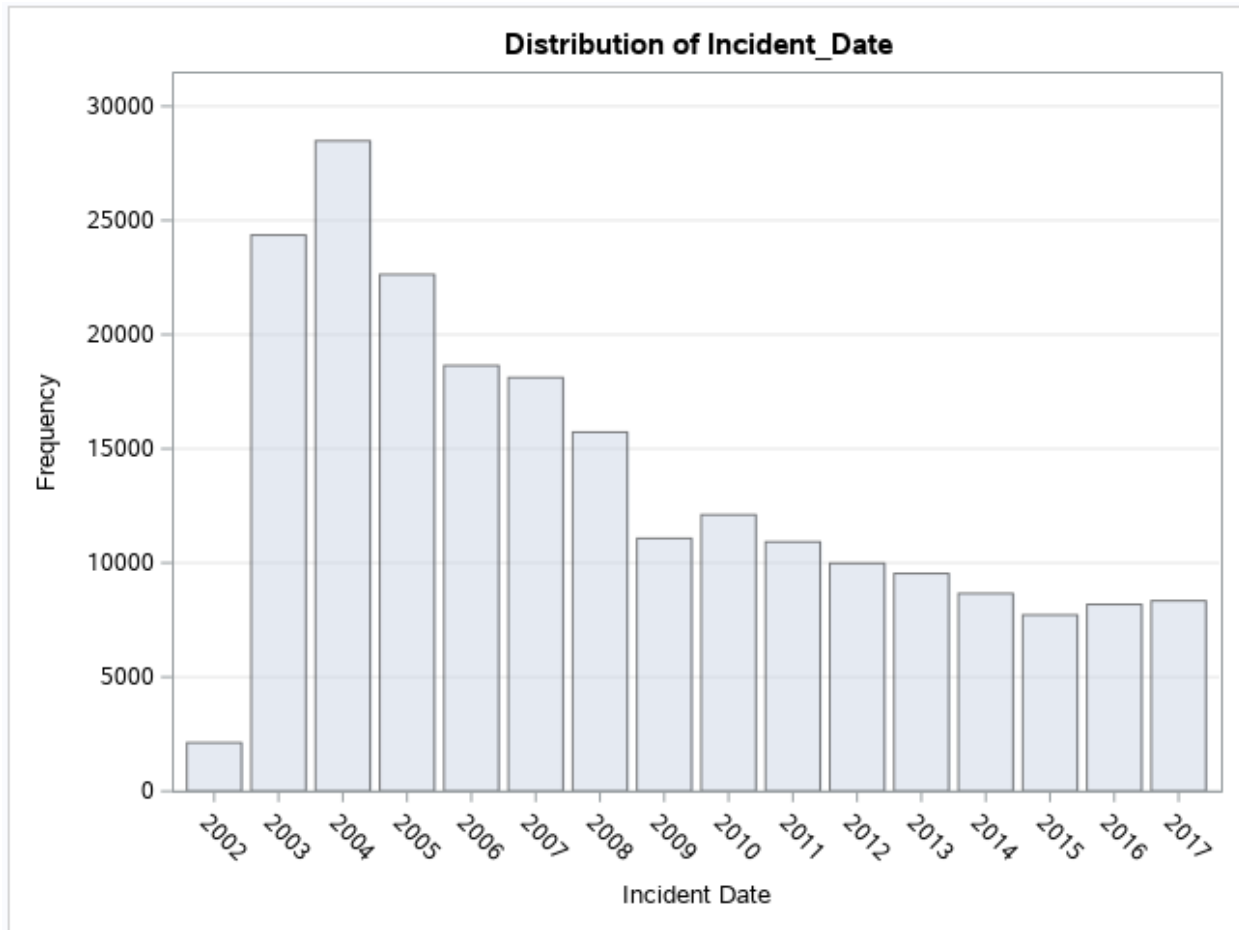
### Overall Data Issues in the Data

Date Issues	
Date_Issues	Frequency
	216609
Needs Review	4241

---

### Overall Claims by Year

Incident Date	
Incident_Date	Frequency
2002	2123
2003	24359
2004	28484
2005	22631
2006	18643
2007	18116
2008	15727
2009	11075
2010	12108
2011	10921
2012	9984
2013	9536
2014	8659
2015	7721
2016	8182
2017	8340



## Hawaii Claim Type, Claim Site, and Disposition

Claim Type	
Claim_Type	Frequency
Passenger Property Loss	2762
Property Damage	1481
Unknown	70
Personal Injury	35
Employee Loss (MPCECA)	13
Passenger Theft	7
Complaint	2
Motor Vehicle	2
Missed Flight	1

Claim Site	
Claim_Site	Frequency
Checked Baggage	3241
Checkpoint	1099
Other	20
Unknown	11
Motor Vehicle	2

Disposition	Frequency
Deny	2030
Approve in Full	1017
Settle	674
Unknown	401
In Review	207
*Insufficient	34
Closed:Canceled	10

### Close Amount Statistics for Hawaii

Analysis Variable : Close_Amount Close Amount			
Mean	Minimum	Maximum	Sum
74	0	5166	279731