

Lab. Work 2

Name: Ravinthiran Partheepan
Course: Data mining Methods
Submission date: 04/10/2023

Problem Statement: Perform Analysis of employee data and airport data.

Tasks:

1. Library reference to your data directory

Code Snippet:

```
libname data '/home/u63603566/Lab-2';
```

Reference:

https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/basess/p0ppzdo5n8sdpdn1mokpbaxt880z.htm

2. Import the emplist data

Employee Data				
Obs	EmployeeID	EmployeeName	Salary	Department
1	.	LANGKAMM	.	LEISTNER
2	.	TOMAS	.	LEHMANN
3	.	TIETZ	.	WALKER
4	.	OESTERBE	.	TORR

Figure 1: Employee Data Points

Code Snippet:

```
data data.emplist;
  infile '/home/u63603566/Lab-2/emplist_data.dat' delimiter=';';
  input emplistID emplistName $ Salary Department $;
run;
```

3. Import the airport data

Airport Data				
Obs	AirportCode	AirportName	City	Country
1	CDGParis	CPHKoben	DFWDalla	TX
2	FCORoma	FRAFrank	GVAGenev	HKGHong
3	HNDTokyo	JFKNew Y	NY	LHRLondo
4	PWMPortl	ME	RDURalei	NC
5	SFOSan F	CA	SYDSydne	New Sout

Figure 2: Airport Data Points

Code Snippet:

```
data data.airport;
  infile '/home/u63603566/Lab-2/airport_data.dat' delimiter=',';
  input AirportCode $ AirportName $ City $ Country $;
run;
```

4. Export emplist data to CSV

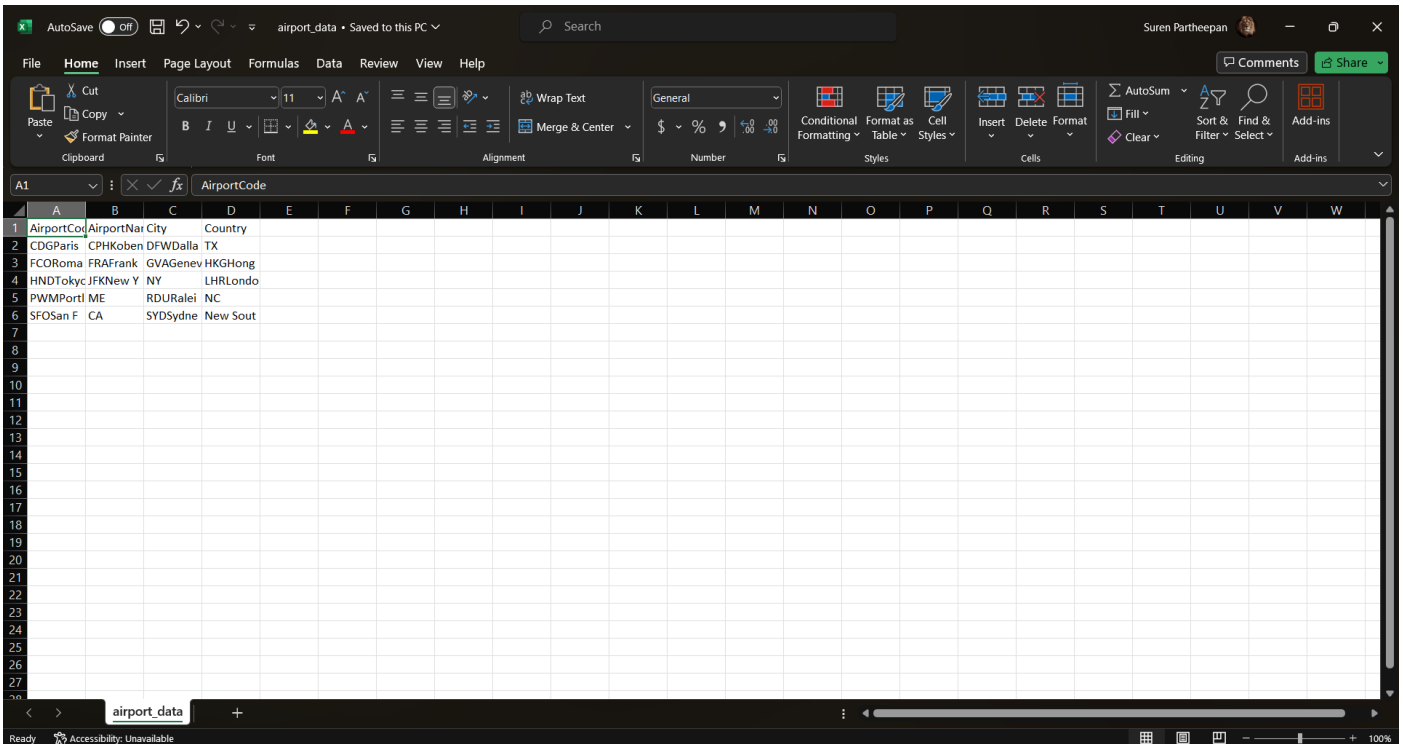
EmployeeID	EmployeeName	Salary	Department
1	LANGKAMM	LEISTNER	
2	TOMAS	LEHMANN	
3	TIETZ	WALKER	
4	OESTERBE	TORR	
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			

Figure 3: Employee Data to CSV Export

Code Snippet:

```
proc export data=data.emplist
  outfile='/home/u63603566/Lab-2/emplist_data.csv'
  dbms=csv replace;
run;
```

5. Export airport data to CSV



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	AirportCode	AirportName	City	Country																			
2	CDG	Paris	CPH	Koblenz	DFW	Dallas	TX																
3	FCO	Roma	FRA	Frankfurt	GVA	Geneva	HKG	Hong Kong															
4	HND	Tokyo	JFK	New York	NY		LHR	London															
5	PWM	Port of Spain	ME		RDU	Raleigh	NC																
6	SFO	San Francisco	CA		SYD	Sydney	New South Wales																
7																							
8																							
9																							
10																							
11																							
12																							
13																							
14																							
15																							
16																							
17																							
18																							
19																							
20																							
21																							
22																							
23																							
24																							
25																							
26																							
27																							
28																							

Figure 4: Airport Data to CSV Export

Code Snippet:

```
proc export data=data.airport
  outfile='/home/u63603566/Lab-2/airport_data.csv'
  dbms=csv replace;
run;
```

6. Explore and view the emplist data

Code Snippet:

```
proc print data=data.emplist;
```

```
title 'emplist Data';
```

```
run;
```

7. Explore and view the airport data

Code Snippet:

```
proc print data=data.airport;
```

```
title 'Airport Data';
```

```
run;
```

8. Perform Frequency analysis on airport data

[2]Cumulative frequencies for Cities and Country features are based on continuous value property.

Frequency of Cities and Countries in Airport Data				
The FREQ Procedure				
City	Frequency	Percent	Cumulative Frequency	Cumulative Percent
DFWDalla	1	20.00	1	20.00
GVAGenev	1	20.00	2	40.00
NY	1	20.00	3	60.00
RDURalei	1	20.00	4	80.00
SYDSydne	1	20.00	5	100.00

Country	Frequency	Percent	Cumulative Frequency	Cumulative Percent
HKGHong	1	20.00	1	20.00
LHRLondo	1	20.00	2	40.00
NC	1	20.00	3	60.00
New Sout	1	20.00	4	80.00
TX	1	20.00	5	100.00

Figure 5: Frequency of Cities and Countries

Code Snippet:

```
proc freq data=data.airport;
```

```
tables City Country;
```

```
title 'Frequency of Cities and Countries in Airport Data';
```

```
run;
```

Reference:

[1]https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/procstat/procstat_freq_synt_ax01.htm

[2]<https://sascrunch.com/proc-freq/> → Refer 5th point in the article

9. Perform Frequency analysis on airport data

[2] Cumulative frequencies for Employee Name and Department features are based on continuous value property.

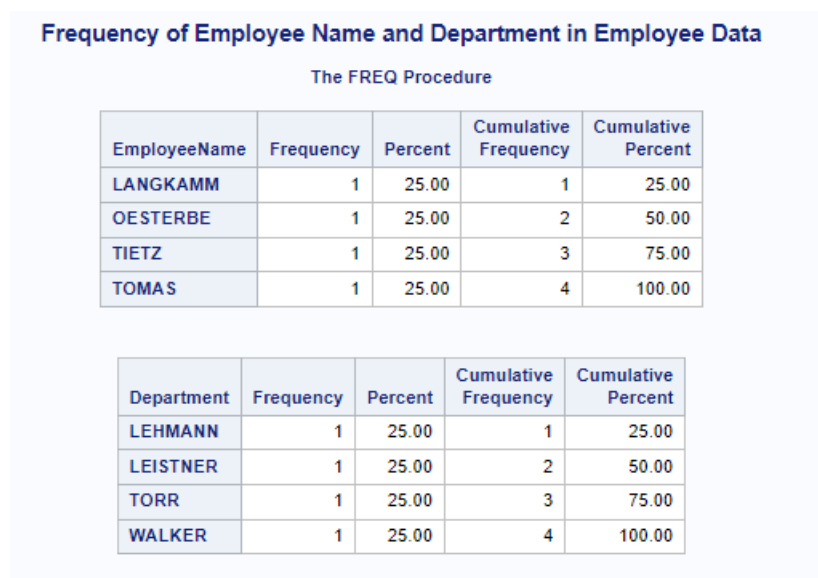


Figure 6: Frequency of Employee Name and Department

Code Snippet:

```
proc freq data=data.airport;
  tables City Country;
  title 'Frequency of Cities and Countries in Airport Data';
run;
```

Reference:

[1]https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/procstat/procstat_freq_synt_ax01.htm

[2]<https://sascrunch.com/proc-freq/> → Refer 5th point in the article

Github Repository: <https://github.com/ravinthiranpartheepan1407/SAS/tree/main/Task-2>