kauno technologijos universitetas
1922

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 from a .dat file**

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 from a .dat file**

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

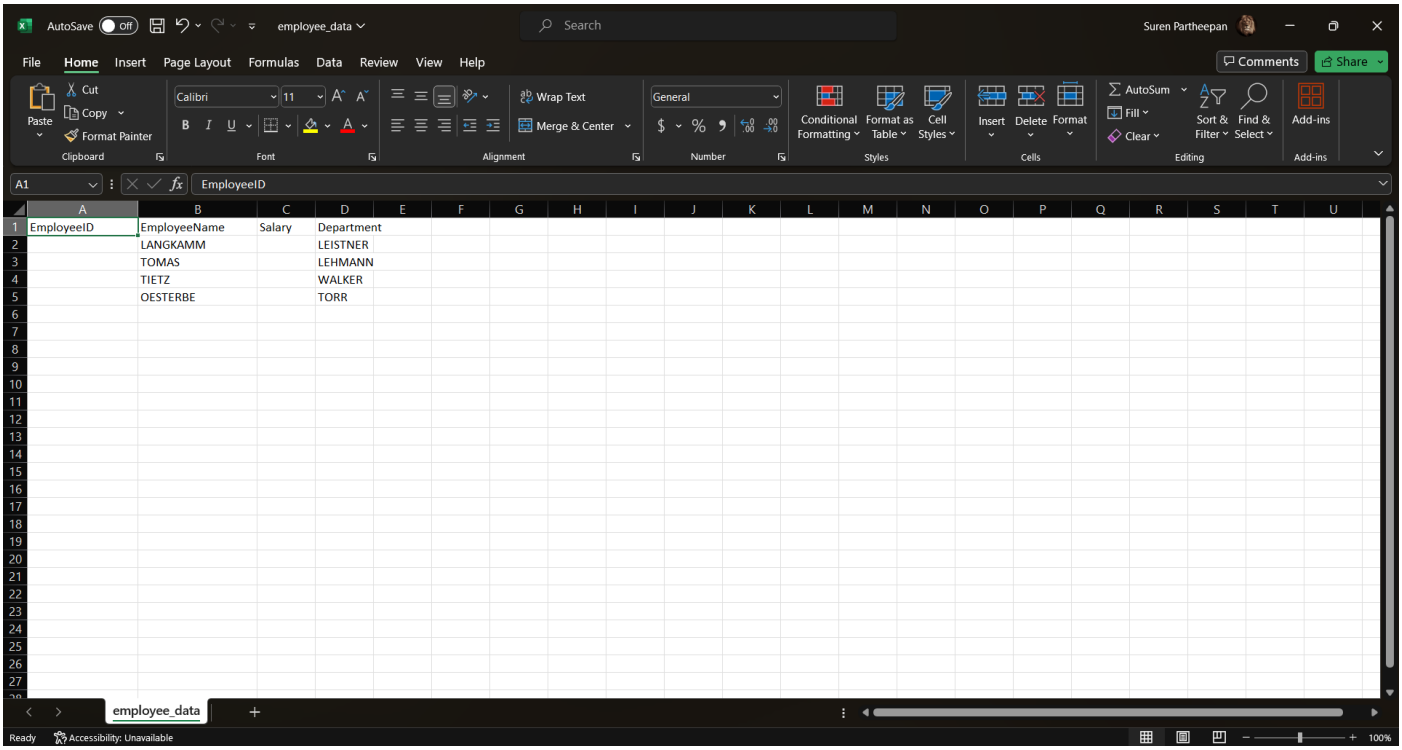

Figure 3: Employee Data to CSV Export

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**

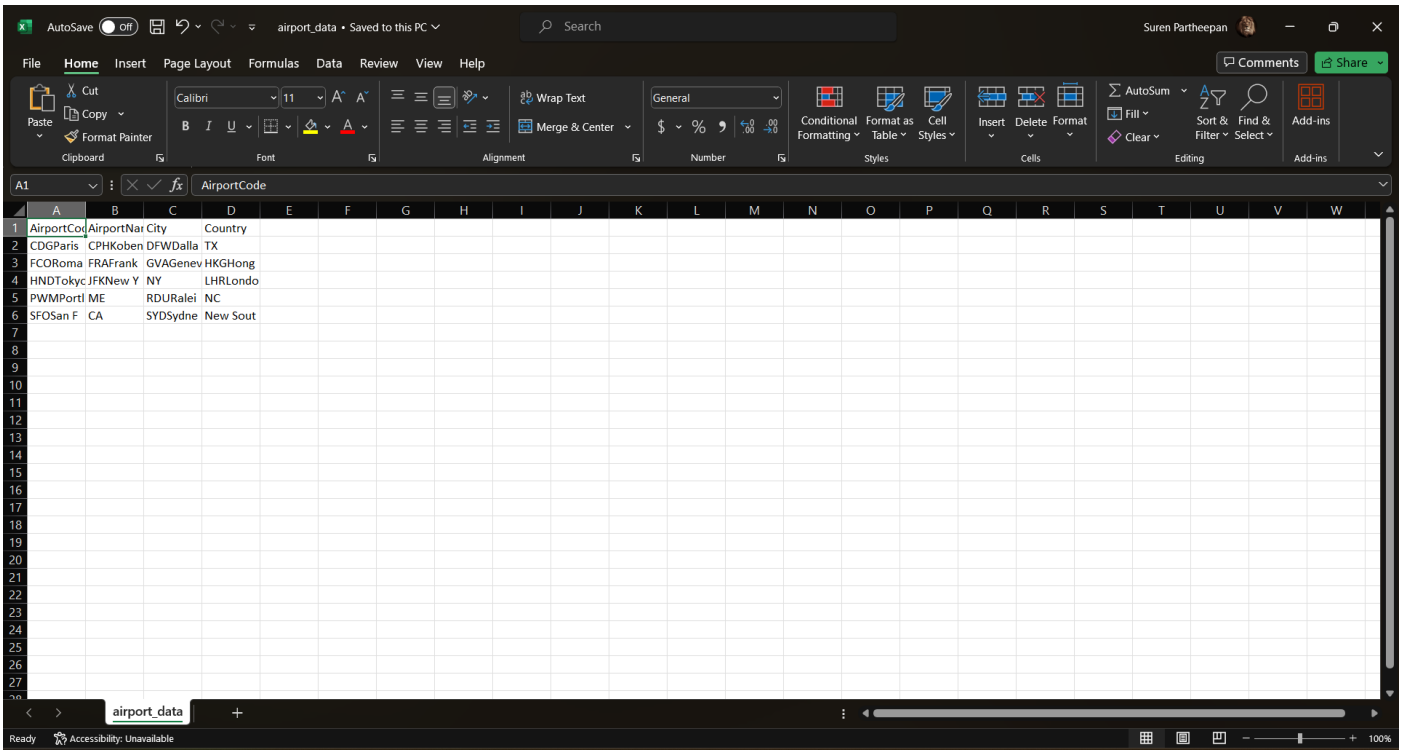

Figure 4: Airport Data to CSV Export

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_syntax01.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.

Frequency of Employee Name and Department in Employee Data

The FREQ Procedure

EmployeeName	Frequency	Percent	Cumulative Frequency	Cumulative Percent
LANGKAMM	1	25.00	1	25.00
OESTERBE	1	25.00	2	50.00
TIETZ	1	25.00	3	75.00
TOMAS	1	25.00	4	100.00

Department	Frequency	Percent	Cumulative Frequency	Cumulative Percent
LEHMANN	1	25.00	1	25.00
LEISTNER	1	25.00	2	50.00
TORR	1	25.00	3	75.00
WALKER	1	25.00	4	100.00

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_syntax01.htm

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

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