A DETAILED ANALYSIS REPORT ON TRAVEL SURVEY OF RESIDENTS OF CANADA 2017.

Moganaviniith Rathinavel(200531668), Nanditha Babu(200573747), Srivas Sathyanarayanan(20054467)

INTRODUCTION

- In this presentation, we delve into a detailed analysis that uncovers valuable insights into the travel patterns, preferences, and behaviors of the residents of this diverse and dynamic nation.

- The 2017 travel survey serves as a crucial lens through which we examine the travel habits that define the Canadian populace. This analysis not only provides a snapshot of the nation's travel landscape but also serves as a valuable resource for stakeholders in the travel and tourism industry, policymakers, and researchers.

- Our dataset encapsulates a myriad of dimensions - from travel durations and preferred destinations to transportation modes and expenditure patterns.

- Through Excel, we've organized and processed this data, extracting meaningful trends and patterns that offer a deep understanding of how Canadians engage with travel.

- Through interactive dashboards and dynamic visualizations, we transform numbers into narratives, providing you with an immersive experience to explore the complex patterns of Canadian travel behavior.

- **Tools used** : SAS studio, excel and Powerbi.

PROJECT TEAM

Moganaviniith Rathinavel	**Team leader**	• **3+ years of experience as a project coordinator in developing business, strategies, signing Mou's with organizations, experience in project management tools like Jira, MS project and monday.com for making perfect schedule.**
Nanditha Babu	**Analyst**	• **3+ Years of experience in collecting data, Data cleaning and preprocessing skills are crucial.** • **Assuring data quality, addressing missing data, and eliminating outliers. which is valuable as a data analyst.** • **A strong understanding of SQL (Structured Query Language) is crucial for data retrieval and manipulation from relational databases.**
Srivas Sathyanarayanan	**Developer**	• **A programmer with more than 2 years of Java, Python, and SQL experience.** • **An expert Python user, web developer, and data analyst. Complex querying, database design, and optimization are all areas of MySQL expertise.** • **Know all stages of the development life cycle, went to client meetings, and collected business analyst's needs.**

PROJECT SUMMARY

- The dataset in question originates from the comprehensive 2017 Travel Survey of Residents of Canada.

- It specifically comprises Public Use Microdata Files (PUMFs) that meticulously detail individual trips undertaken by Canadian residents throughout the specified year.

- This dataset serves as a valuable resource for researchers, policymakers, and analysts, offering a granular perspective on travel patterns, preferences, and behaviors within the Canadian population during the surveyed period.

The data dictionary outlines various variables, their positions, descriptions, and corresponding pages in the dataset. Here are some key points:

Reference Information:

- REFYEAR and REFMONTH represent the reference year and month. PUMFID and TRIPID are identifiers for the public use microdata file and trip, respectively. QUARTER indicates the quarter of the reference period.

Trip Origin and Destination:

- ORCPROVT, ORCCDT2, and ORCCMAT2 provide information about the province, census division, and census metropolitan area of trip origin. MDDPLFL, MDCCD, and MDCCMA2 offer similar details for the trip destination.

Demographic and Economic Factors:

- AGE_GR2, SEX, EDLEVGR, LFSSTATG, and INCOMGR2 represent demographic and economic characteristics of the respondent.

Trip Details:

- Various variables (e.g., TMDTYPE2, DIST2, MRDTRIP2, MRDTRIP3) describe trip-related information such as transportation mode, trip distance, and main reason for the trip.

Accommodation and Nights Spent:

- Information about the nights spent in different provinces and territories, types of accommodations, and various spending categories are provided.

Spending Details:

- Total spending and spending in different categories, including accommodations, vehicle rental, food, and recreational activities, are outlined.

Package Information:

- Package-related details, including spending allocation and flags indicating package components, are included.

Additional Activities:

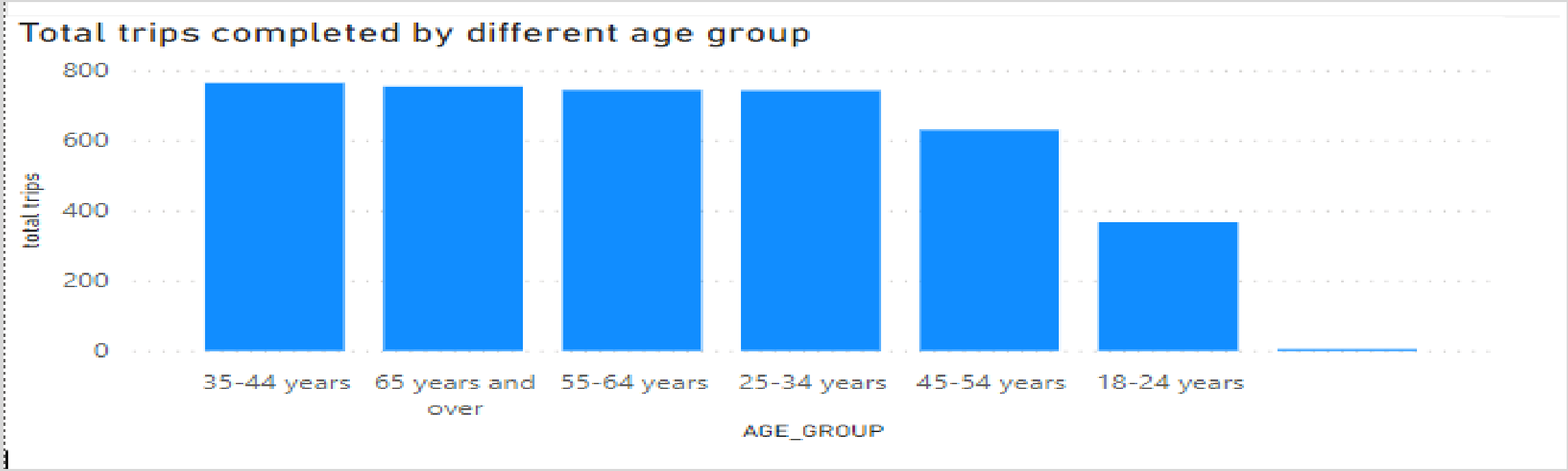
- The dataset contains information about various activities undertaken during the trip, such as attending events, going to the beach, boating, and more.

THE KEYVARIABLES USED AND FOUR RESEARCH QUESTIONS

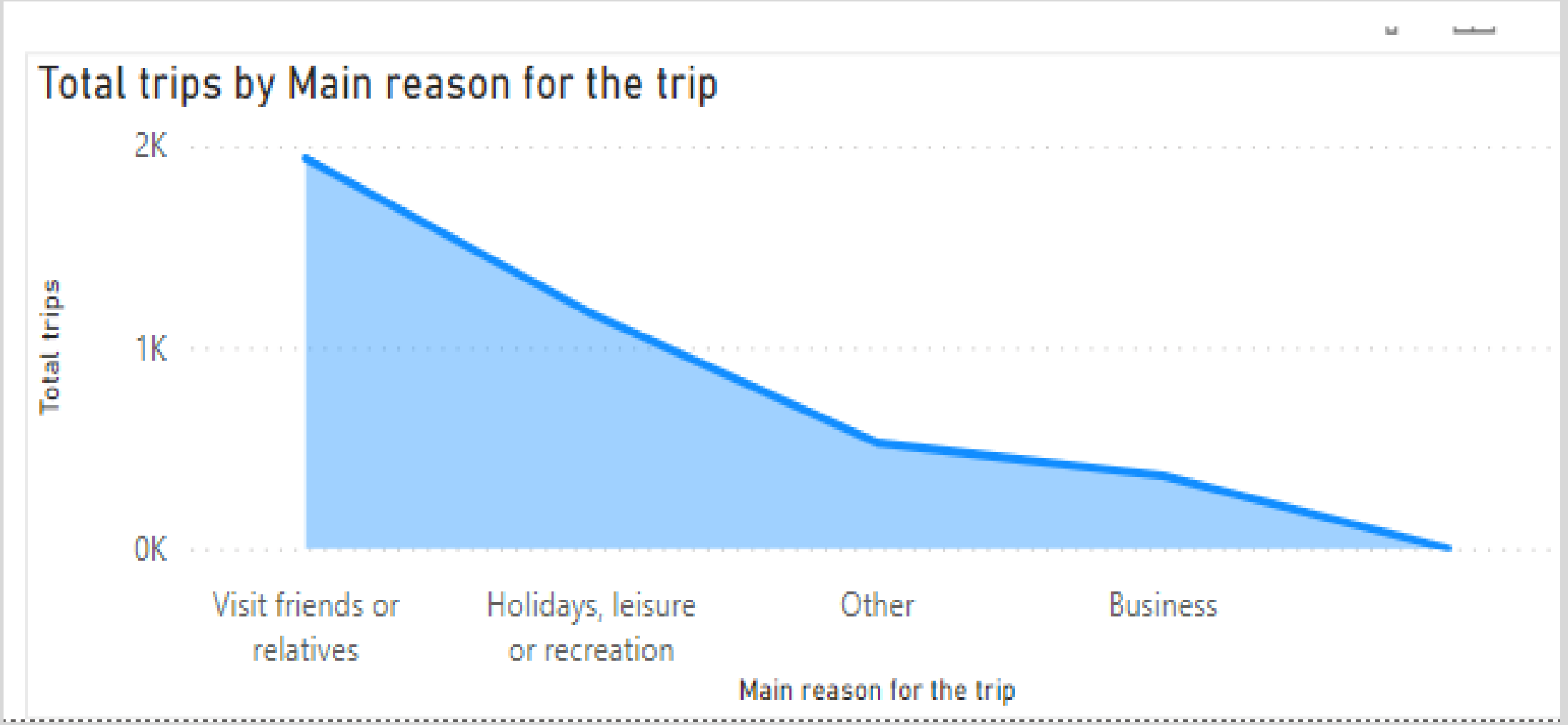
- The key variables used from the data are AGE GROUP, SEX, Province of trip Destination and trip type with respect to total trips.

FOUR RESEARCH QUESTIONS:

1.What is the total number of trips generated by different age groups?



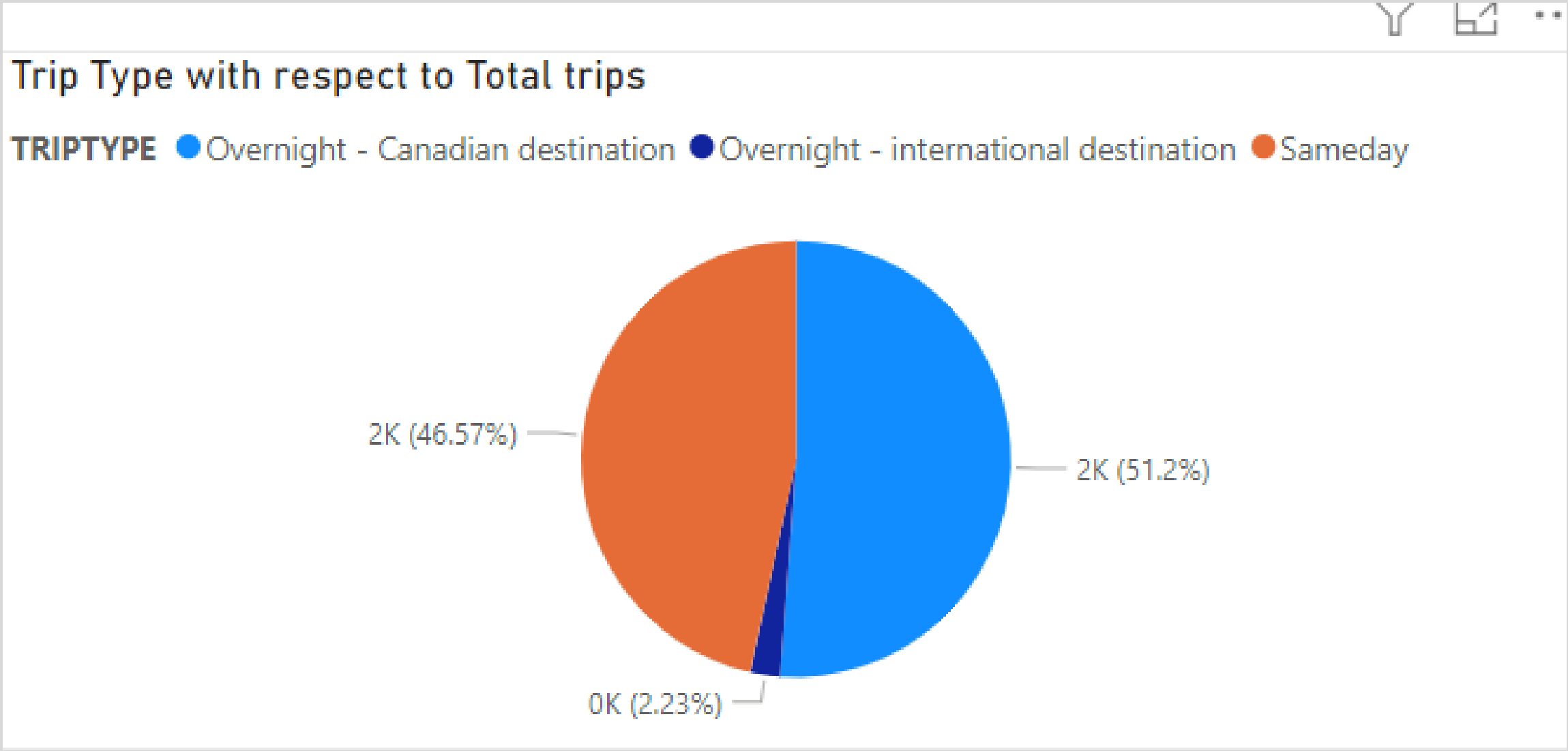
2.What is the main reason for the trip?



3.Which provinces in Canada were visited by visitors?

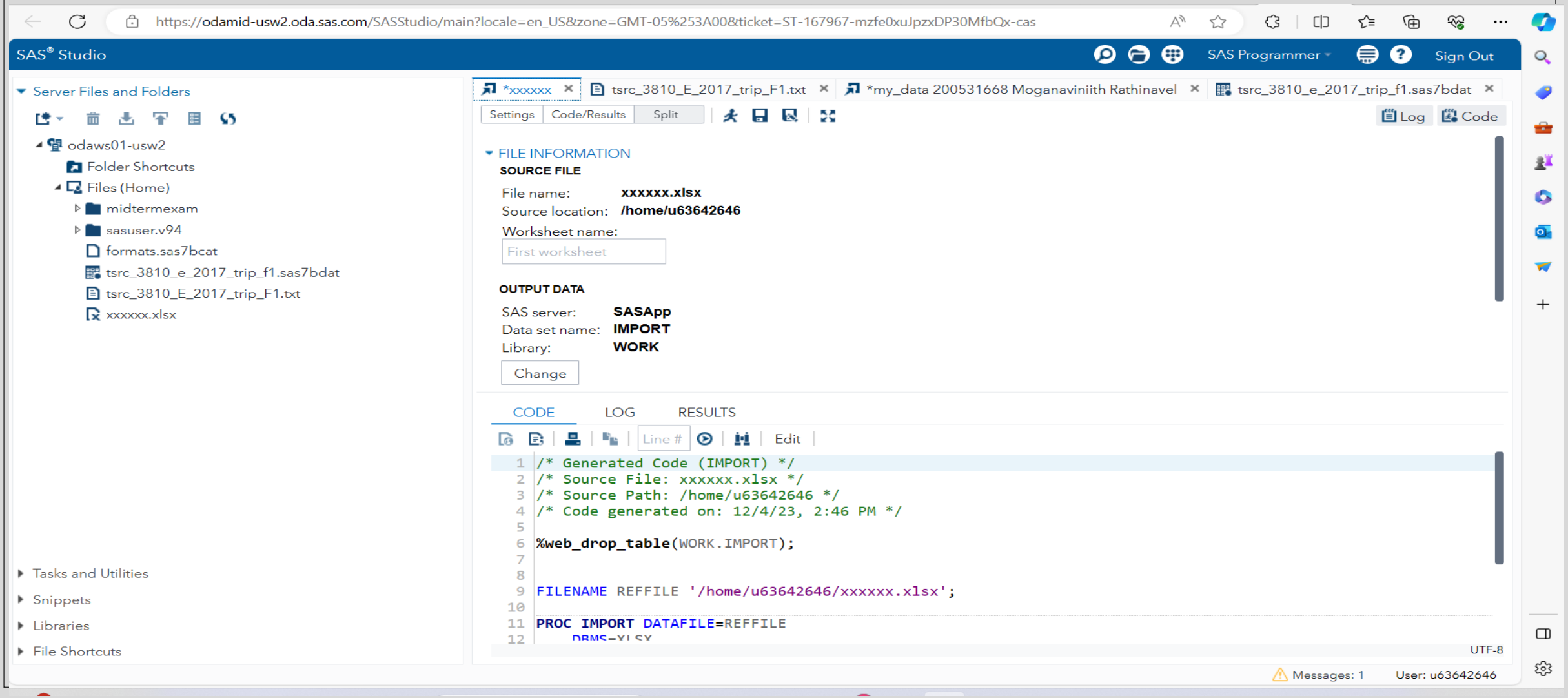
Province of trip Destination		Alberta	British Columbia	Manitoba	New Brunswick	Newfoundlar
	1					
Alberta		366	30	13	1	
British Columbia		42	312	8	1	
Manitoba		3	1	233	1	
New Brunswick		2	1	1	157	
Newfoundland and Labrador		3			1	
Northwest Territories			1			
Not stated			1			
Nova Scotia					23	
Ontario		6	11	26	6	
Prince Edward Island					4	
Quebec		3	1	3	13	
Saskatchewan		14	3	17		
Yukon		1				
Total	**1**	**440**	**361**	**301**	**207**	

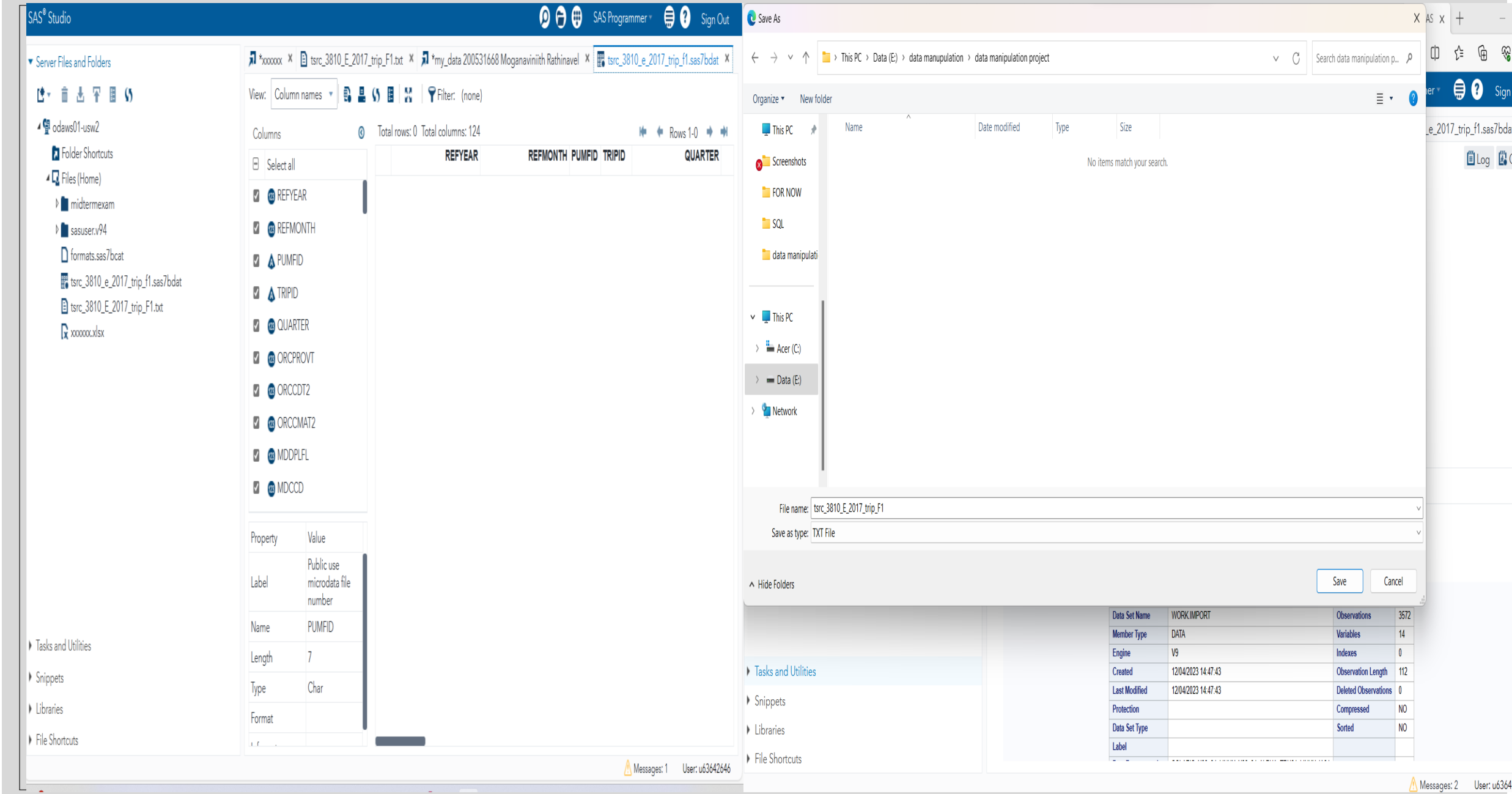
4.What is the most preferred trip type among visitors?



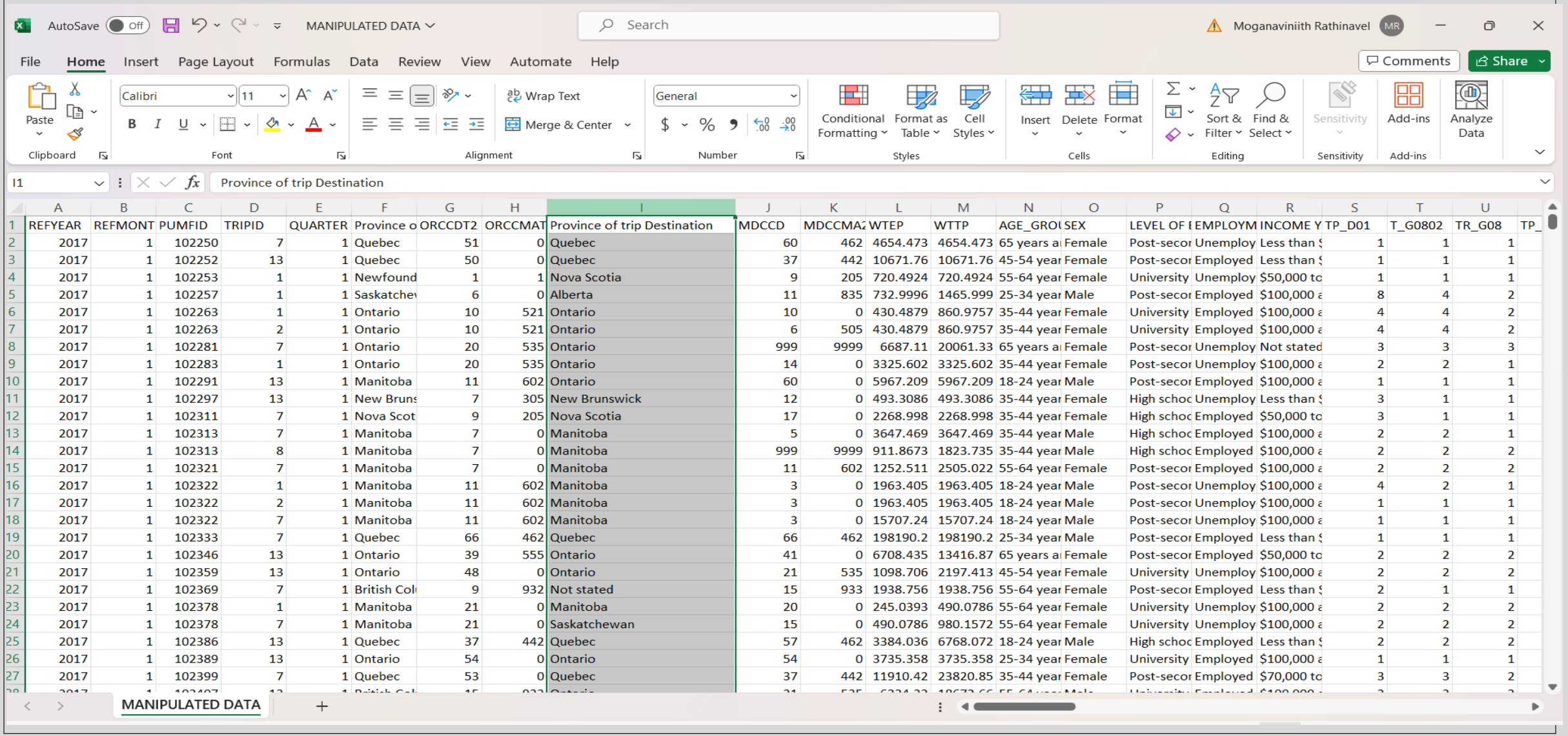
DATA MANIPULATION STEPS

EXPORTING DATA FROM SAS STUDIO




SAS® Studio
SAS Programmer
Sign Out
Server Files and Folders
odaws01-usw2
Folder Shortcuts
Files (Home)
midtermexam
sasuser.v94
formats.sas7bcat
tsrc_3810_e_2017_trip_f1.sas7bdat
tsrc_3810_E_2017_trip_F1.txt
xxxxxx.xlsx
*xxxxxx
tsrc_3810_E_2017_trip_F1.txt
*my_data 200531668 Moganaviniith Rathinavel
tsrc_3810_e_2017_trip_f1.sas7bdat
View: Column names
Filter: (none)
Columns
Total rows: 0 Total columns: 124
Rows 1-0
REFYEAR REFMONTH PUMFID TRIPID QUARTER
Select all
REFYEAR
REFMONTH
PUMFID
TRIPID
QUARTER
ORCPROVT
ORCCDT2
ORCCMAT2
MDDPLFL
MDCCD
Property
Value
Label
Public use microdata file number
Name
PUMFID
Length
7
Type
Char
Format
Tasks and Utilities
Snippets
Libraries
File Shortcuts
Messages: 1
User: u63642646
Save As
This PC › Data (E:) › data manupulation › data manipulation project
Search data manipulation p...
Organize
New folder
This PC
Screenshots
FOR NOW
SQL
data manipulati
Name
Date modified
Type
Size
No items match your search.
This PC
Acer (C:)
Data (E:)
Network
File name: tsrc_3810_E_2017_trip_F1
Save as type: TXT File
Hide Folders
Save
Cancel
Log
Code
Data Set Name WORK.IMPORT Observations 3572
Member Type DATA Variables 14
Engine V9 Indexes 0
Created 12/04/2023 14:47:43 Observation Length 112
Last Modified 12/04/2023 14:47:43 Deleted Observations 0
Protection Compressed NO
Data Set Type Sorted NO
Label
Messages: 2
User: u63642646

CLEANING AND MANIPULATING THE DATA(CSV FILE)USING EXCEL



AutoSave Off | MANIPULATED DATA | Search | Moganaviniith Rathinavel MR

File Home Insert Page Layout Formulas Data Review View Automate Help | Comments | Share

Paste | Calibri 11 | Wrap Text | Merge & Center | General | Conditional Formatting | Format as Table | Cell Styles | Insert | Delete | Format | Sort & Filter | Find & Select | Sensitivity | Add-ins | Analyze Data

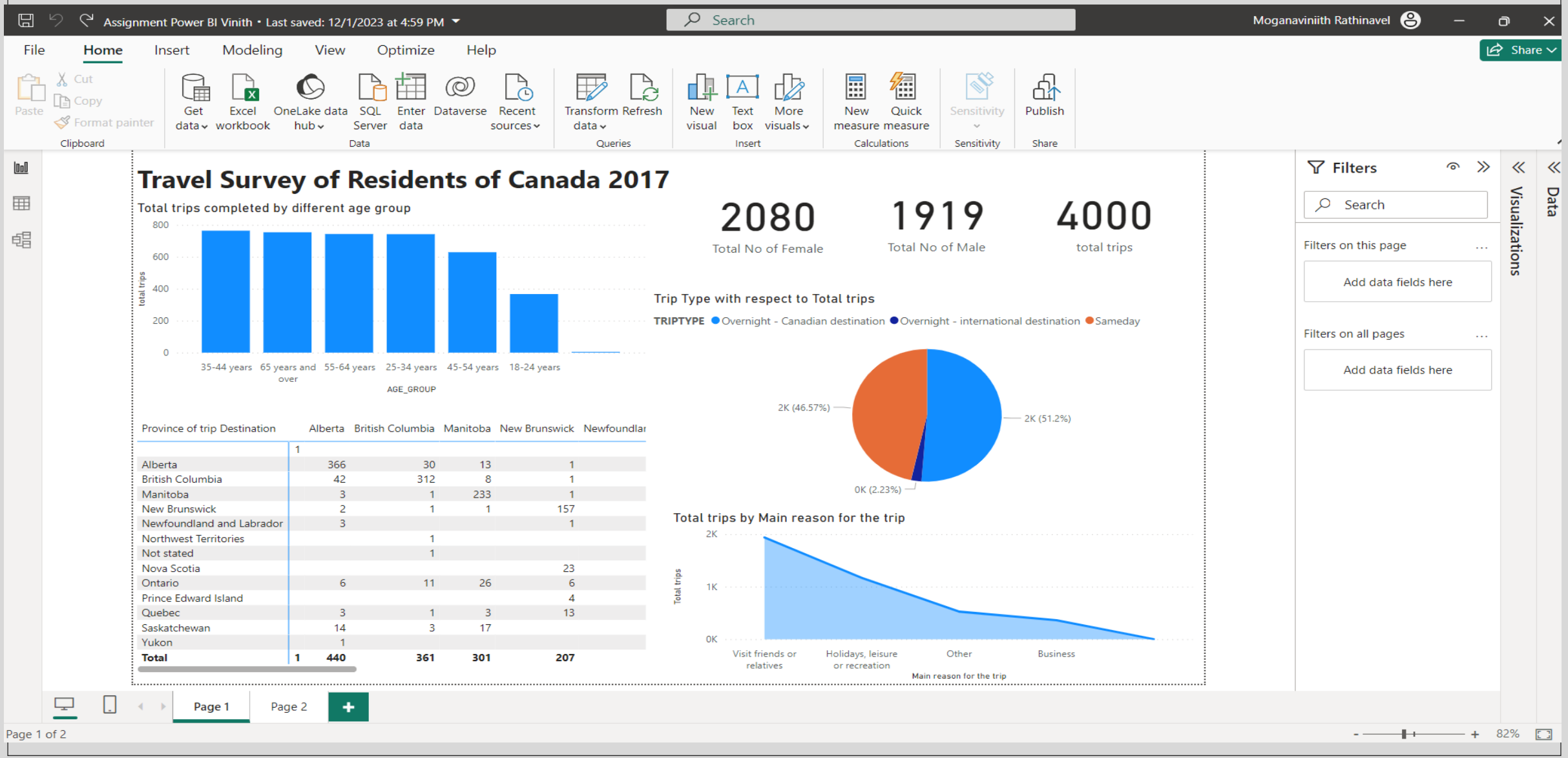
Clipboard | Font | Alignment | Number | Styles | Cells | Editing | Sensitivity | Add-ins

I1 | Province of trip Destination

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	REFYEAR	REFMONT	PUMFID	TRIPID	QUARTER	Province o	ORCCDT2	ORCCMAT	Province of trip Destination	MDCCD	MDCCMA2	WTEP	WTTP	AGE_GROU	SEX	LEVEL OF E	EMPLOYM	INCOME Y	TP_D01	T_G0802	TR_G08	TP_
2	2017	1	102250	7	1	Quebec	51	0	Quebec	60	462	4654.473	4654.473	65 years a	Female	Post-secor	Unemploy	Less than $	1	1	1	
3	2017	1	102252	13	1	Quebec	50	0	Quebec	37	442	10671.76	10671.76	45-54 year	Female	Post-secor	Employed	Less than $	1	1	1	
4	2017	1	102253	1	1	Newfound	1	1	Nova Scotia	9	205	720.4924	720.4924	55-64 year	Female	University	Unemploy	$50,000 to	1	1	1	
5	2017	1	102257	1	1	Saskatchev	6	0	Alberta	11	835	732.9996	1465.999	25-34 year	Male	Post-secor	Employed	$100,000 a	8	4	2	
6	2017	1	102263	1	1	Ontario	10	521	Ontario	10	0	430.4879	860.9757	35-44 year	Female	University	Employed	$100,000 a	4	4	2	
7	2017	1	102263	2	1	Ontario	10	521	Ontario	6	505	430.4879	860.9757	35-44 year	Female	University	Employed	$100,000 a	4	4	2	
8	2017	1	102281	7	1	Ontario	20	535	Ontario	999	9999	6687.11	20061.33	65 years a	Female	Post-secor	Unemploy	Not stated	3	3	3	
9	2017	1	102283	1	1	Ontario	20	535	Ontario	14	0	3325.602	3325.602	35-44 year	Female	Post-secor	Unemploy	$100,000 a	2	2	1	
10	2017	1	102291	13	1	Manitoba	11	602	Ontario	60	0	5967.209	5967.209	18-24 year	Male	Post-secor	Employed	$100,000 a	1	1	1	
11	2017	1	102297	13	1	New Bruns	7	305	New Brunswick	12	0	493.3086	493.3086	35-44 year	Female	High schoc	Unemploy	Less than $	3	1	1	
12	2017	1	102311	7	1	Nova Scot	9	205	Nova Scotia	17	0	2268.998	2268.998	35-44 year	Female	High schoc	Employed	$50,000 to	3	1	1	
13	2017	1	102313	7	1	Manitoba	7	0	Manitoba	5	0	3647.469	3647.469	35-44 year	Male	High schoc	Employed	$100,000 a	2	2	1	
14	2017	1	102313	8	1	Manitoba	7	0	Manitoba	999	9999	911.8673	1823.735	35-44 year	Male	High schoc	Employed	$100,000 a	2	2	2	
15	2017	1	102321	7	1	Manitoba	7	0	Manitoba	11	602	1252.511	2505.022	55-64 year	Female	Post-secor	Employed	$100,000 a	2	2	2	
16	2017	1	102322	1	1	Manitoba	11	602	Manitoba	3	0	1963.405	1963.405	18-24 year	Male	Post-secor	Unemploy	$100,000 a	4	2	1	
17	2017	1	102322	2	1	Manitoba	11	602	Manitoba	3	0	1963.405	1963.405	18-24 year	Male	Post-secor	Unemploy	$100,000 a	1	1	1	
18	2017	1	102322	7	1	Manitoba	11	602	Manitoba	3	0	15707.24	15707.24	18-24 year	Male	Post-secor	Unemploy	$100,000 a	1	1	1	
19	2017	1	102333	7	1	Quebec	66	462	Quebec	66	462	198190.2	198190.2	25-34 year	Male	Post-secor	Employed	Less than $	1	1	1	
20	2017	1	102346	13	1	Ontario	39	555	Ontario	41	0	6708.435	13416.87	65 years a	Female	Post-secor	Employed	$50,000 to	2	2	2	
21	2017	1	102359	13	1	Ontario	48	0	Ontario	21	535	1098.706	2197.413	45-54 year	Female	University	Unemploy	$100,000 a	2	2	2	
22	2017	1	102369	7	1	British Col	9	932	Not stated	15	933	1938.756	1938.756	55-64 year	Female	Post-secor	Employed	Less than $	2	1	1	
23	2017	1	102378	1	1	Manitoba	21	0	Manitoba	20	0	245.0393	490.0786	55-64 year	Female	University	Unemploy	$100,000 a	2	2	2	
24	2017	1	102378	7	1	Manitoba	21	0	Saskatchewan	15	0	490.0786	980.1572	55-64 year	Female	University	Unemploy	$100,000 a	2	2	2	
25	2017	1	102386	13	1	Quebec	37	442	Quebec	57	462	3384.036	6768.072	18-24 year	Male	High schoc	Employed	Less than $	2	2	2	
26	2017	1	102389	13	1	Ontario	54	0	Ontario	54	0	3735.358	3735.358	25-34 year	Female	University	Employed	$100,000 a	1	1	1	
27	2017	1	102399	7	1	Quebec	53	0	Quebec	37	442	11910.42	23820.85	35-44 year	Female	Post-secor	Employed	$70,000 to	3	3	2	

MANIPULATED DATA

DATA VISULIZATION USING POWER BI



Province of trip Destination	Alberta	British Columbia	Manitoba	New Brunswick	Newfoundlar
	1				
Alberta	366	30	13	1	
British Columbia	42	312	8	1	
Manitoba	3	1	233	1	
New Brunswick	2	1	1	157	
Newfoundland and Labrador	3			1	
Northwest Territories		1			
Not stated		1			
Nova Scotia				23	
Ontario	6	11	26	6	
Prince Edward Island				4	
Quebec	3	1	3	13	
Saskatchewan	14	3	17		
Yukon	1				
Total	**1 440**	**361**	**301**	**207**	

Here are some common problems that were raised during data preparation:

Missing Data:

- Incomplete or missing values can be a common issue. Deciding how to handle missing data, whether through imputation or exclusion, requires careful consideration.

Data Cleaning:

- Inconsistent or inaccurate entries may need to be identified and corrected. This includes handling outliers and anomalies that could skew the analysis.

Data Integration:

- Combining data from multiple sources can present challenges in terms of aligning formats, resolving discrepancies, and ensuring consistency.

Normalization and Standardization:

- Variables might need to be normalized or standardized to ensure that they are on a consistent scale, especially when different measurement units are involved.

Dealing with Categorical Data:

- Converting categorical variables into a format suitable for analysis, such as one-hot encoding or label encoding, can be challenging.

Ensuring Data Quality:

- Verifying the quality of the data is crucial. This involves checking for data accuracy, consistency, and reliability.

Handling Duplicates:

- Duplicate entries, if present, need to be identified and addressed to avoid duplication of information in the analysis.

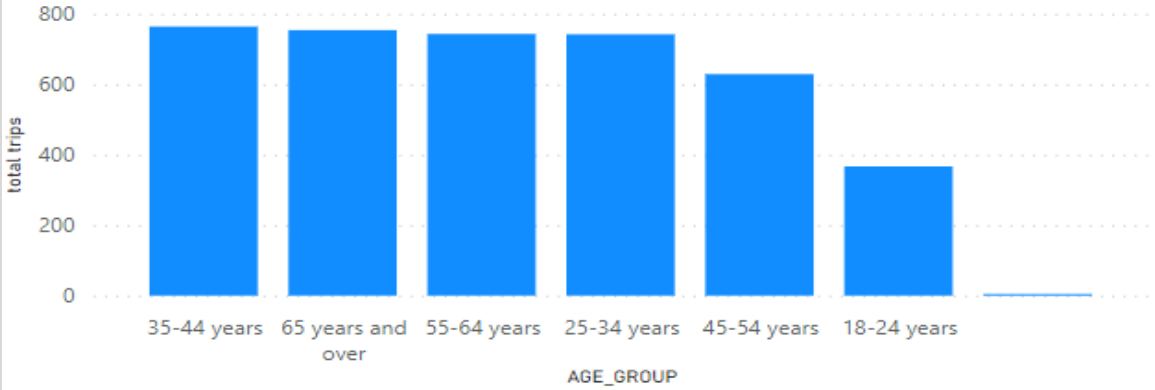
Data Security and Privacy:

- Ensuring that sensitive information is handled securely, and that the dataset complies with privacy regulations is a critical consideration.

POWERBI DASHBOARD

Travel Survey of Residents of Canada 2017

Total trips completed by different age group



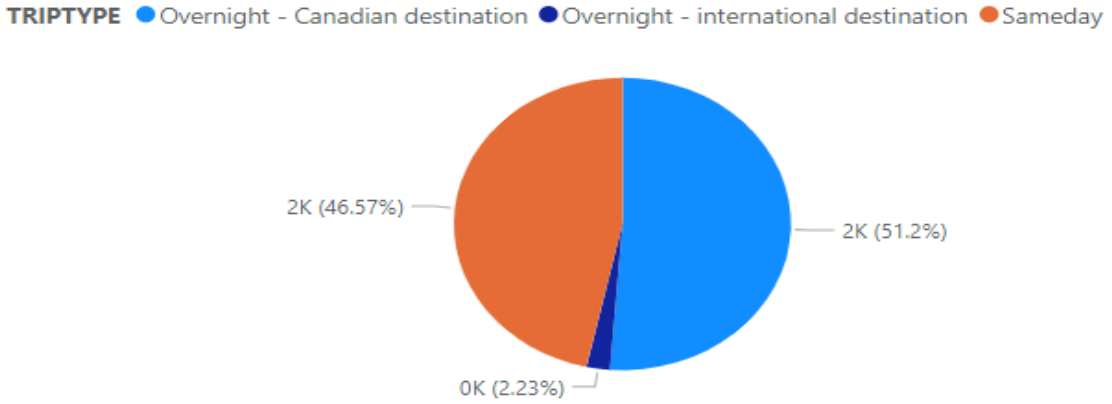
Province of trip Destination	Alberta	British Columbia	Manitoba	New Brunswick	Newfoundlar
	1				
Alberta	366	30	13	1	
British Columbia	42	312	8	1	
Manitoba	3	1	233	1	
New Brunswick	2	1	1	157	
Newfoundland and Labrador	3			1	
Northwest Territories		1			
Not stated		1			
Nova Scotia				23	
Ontario	6	11	26	6	
Prince Edward Island				4	
Quebec	3	1	3	13	
Saskatchewan	14	3	17		
Yukon	1				
Total	**1 440**	**361**	**301**	**207**	

2080
Total No of Female

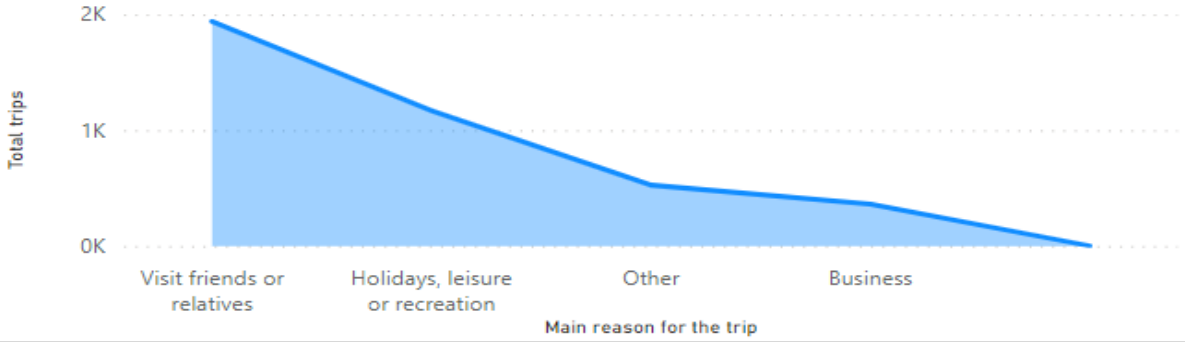
1919
Total No of Male

4000
total trips

Trip Type with respect to Total trips



Total trips by Main reason for the trip



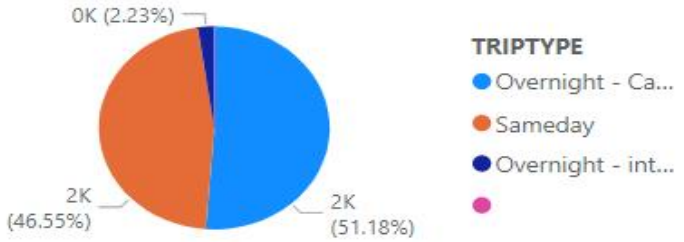
POWERBI DASHBOARD

2080
Total No of Female

1919
Total No of Male

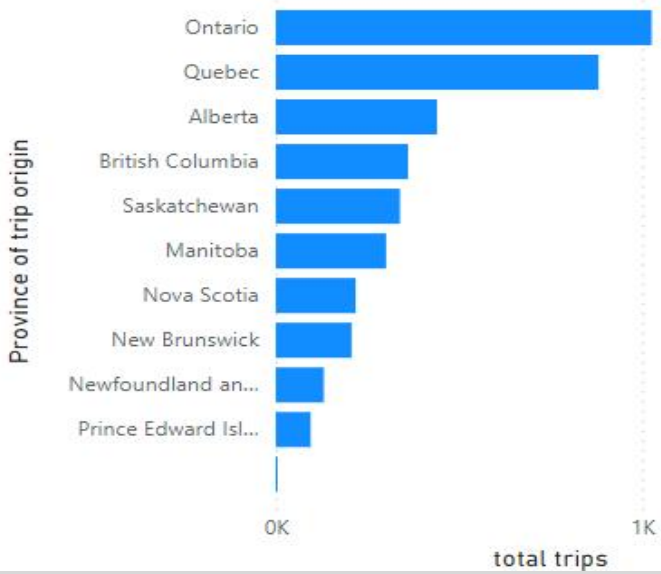
4000
total trips

total trips by TRIPTYPE



total trips by Main reason for the trip

total trips by Province of trip origin



REFERENCES:

- Travel Survey of Residents of Canada, 2017: Trip File (oclc.org)
- Create a Power BI dashboard from a report - Power BI | Microsoft Learn


Thank You