

TEAM PROJECT ASSIGNMENT 1

CINS 370: Group 1

Mansi Dogra

Viraj Sonavane

Adhaar Jaiswal

The datasets that are shown below are composed of multiple files that are related to each other. These datasets are an attempt to evaluate large volumes of stored data and are important to support new information that can be derived from the stored data. Each data set is the unit to measure the information released in a public open data repository.

The datasets have several attributes that shows the different properties shared among different objects or activities.

Table 1: Income

Country Code - It is the short form of different country names.

Region - It describes the different parts of the world.

Income Group - It describes upper and lower level of country income.

Country - It is the list of nations on earth.

Reference : <https://www.census.gov/data/tables/2019/demo/income-poverty/p60-266.html>

Table 2 : Product

Region - It describes the different parts of the world.

Country - It is the list of nations on earth.

Item Type - It describes the highest ordered in 2018 by that country.

Reference : <http://eforexcel.com/wp/downloads-18-sample-csv-files-data-sets-for-testing-sales/>

Table 3 : Immigration

Port Name - It describes the ship ports in that country.

State - It describes the name of states in specific countries.

Port Code - It describes specific zip code of the ports.

Border - It describes number of inter related crossings between the countries.

Date - It describes the date of immigration.

Container - It lists number of vehicles migrating via border.

Value - It describes the number of people crossing the border.

Location - It describes the latitude and longitude of the border.

Reference : <https://www.kaggle.com/akhilv11/border-crossing-entry-data>

Table 4 : Population

State - It describes the name of states in specific countries.

Population - It describes the number of US population in 2018.

Reference : <https://www.kaggle.com/lucasvictor/us-state-populations-2018>

Table 5 : Poverty

State - It describes the name of states in specific countries.

People in Poverty - It describes the number of people under or over the poverty line.

Children in Poverty - It describes the number of children under or over the poverty line.

Reference : <https://data.ers.usda.gov/reports.aspx?ID=17826>

1. Income Table

1	Country Code	Region	Income Group	Country
2	ABW	Latin America & Caribbean	High income	Aruba
3	AFG	South Asia	Low income	Afghanistan
4	AGO	Sub-Saharan Africa	Upper middle income	Angola
5	ALB	Europe & Central Asia	Upper middle income	Albania
6	AND	Europe & Central Asia	High income	Andorra
7	ARE	Middle East & North Africa	High income	United Arab Emirates
8	ARG	Latin America & Caribbean	Upper middle income	Argentina
9	ARM	Europe & Central Asia	Lower middle income	Armenia
10	ASM	East Asia & Pacific	Upper middle income	American Samoa
11	ATG	Latin America & Caribbean	High income	Antigua and Barbuda
12	AUS	East Asia & Pacific	High income	Australia
13	AUT	Europe & Central Asia	High income	Austria
14	AZE	Europe & Central Asia	Upper middle income	Azerbaijan
15	BDI	Sub-Saharan Africa	Low income	Burundi
16	BEL	Europe & Central Asia	High income	Belgium
17	BEN	Sub-Saharan Africa	Low income	Benin
18	BFA	Sub-Saharan Africa	Low income	Burkina Faso
19	BGD	South Asia	Lower middle income	Bangladesh
20	BGR	Europe & Central Asia	Upper middle income	Bulgaria

3. Immigration Table

1	Port Name	State	Port Code	Border	Date	Container	Value	Location
2	Calixto East	California	2507	US-Mexico Border	3/2/19 0:00	Trucks	34447	POINT (-115.48433000000001 32.67524)
3	Van Buren	Maine	108	US-Canada Border	3/2/19 0:00	Rail Containers Full	428	POINT (-67.94271 47.16207)
4	Otay Mesa	California	2506	US-Mexico Border	3/2/19 0:00	Trucks	81217	POINT (-117.05333 32.57333)
5	Nogales	Arizona	2604	US-Mexico Border	3/2/19 0:00	Trains	62	POINT (-110.93361 31.340279999999996)
6	Trout River	New York	715	US-Canada Border	3/2/19 0:00	Personal Vehicle Passengers	16377	POINT (-73.44253 44.9900100000000005)
7	Madawaska	Maine	109	US-Canada Border	3/2/19 0:00	Trucks	179	POINT (-68.3271 47.35446)
8	Pembina	North Dakota	3401	US-Canada Border	3/2/19 0:00	Bus Passengers	1054	POINT (-97.24333 48.96639)
9	Progreso	Texas	2309	US-Mexico Border	3/2/19 0:00	Truck Containers Empty	1808	POINT (-97.94889 26.0616700000000003)
10	Portal	North Dakota	3403	US-Canada Border	3/2/19 0:00	Rail Containers Empty	6685	POINT (-102.54917 48.99583)
11	Champlain-Rouses Point	New York	712	US-Canada Border	3/2/19 0:00	Trucks	24759	POINT (-73.44694 44.98639)
12	Ophelm	Montana	3317	US-Canada Border	3/2/19 0:00	Personal Vehicles	235	POINT (-106.40265 48.85574)
13	Neche	North Dakota	3404	US-Canada Border	3/2/19 0:00	Trucks	584	POINT (-97.55444 48.98611)
14	Lancaster	Minnesota	3430	US-Canada Border	3/2/19 0:00	Personal Vehicles	1612	POINT (-96.82024 48.8491100000000001)
15	Derby Line	Vermont	209	US-Canada Border	3/2/19 0:00	Buses	80	POINT (-72.09944 45.005)
16	Sarles	North Dakota	3409	US-Canada Border	3/2/19 0:00	Personal Vehicle Passengers	509	POINT (-98.99457 48.9410500000000004)
17	Wildhorse	Montana	3323	US-Canada Border	3/2/19 0:00	Truck Containers Full	122	POINT (-109.67761 48.54863)
18	Derby Line	Vermont	209	US-Canada Border	3/2/19 0:00	Truck Containers Full	6483	POINT (-72.09944 45.005)
19	Lynden	Washington	3023	US-Canada Border	3/2/19 0:00	Pedestrians	79	POINT (-122.44316000000002 48.94802)
20	Vanceboro	Maine	105	US-Canada Border	3/2/19 0:00	Pedestrians	3	POINT (-67.42955 45.55984)

2. Product Table

1	Region	Country	Item Type
2	Middle East and North Africa	Libya	Cosmetics
3	North America	Canada	Vegetables
4	Middle East and North Africa	Libya	Baby Food
5	Asia	Japan	Cereal
6	Sub-Saharan Africa	Chad	Fruits
7	Europe	Armenia	Cereal
8	Sub-Saharan Africa	Eritrea	Cereal
9	Europe	Montenegro	Clothes
10	Central America and the Caribbean	Jamaica	Vegetables
11	Australia and Oceania	Fiji	Clothes
12	Sub-Saharan Africa	Togo	Snacks
13	Europe	Montenegro	Household
14	Europe	Greece	Cosmetics
15	Sub-Saharan Africa	Sudan	Fruits
16	Asia	Maldives	Clothes
17	Europe	Montenegro	Office Supplies
18	Europe	Estonia	Beverages
19	North America	Greenland	Clothes
20	Sub-Saharan Africa	Cape Verde	Clothes

4. Population Table

COUNTRY USA	2018 Population
California	39776830
Texas	28704330
Florida	21312211
New York	19862512
Pennsylvania	12823989
Illinois	12768320
Ohio	11694664
Georgia	10545138
North Carolina	10390149
Michigan	9991177
New Jersey	9032872
Virginia	8525660
Washington	7530552
Arizona	7123898
Massachusetts	6895917
Tennessee	6782564
Indiana	6699629
Missouri	6135888

5. Poverty Table

USA	People in Poverty	Children in Poverty
State	Lower Level	Lower Bound
Alabama	784,317	253,694
Alaska	75,273	23,957
Arizona	1,002,773	324,723
Arkansas	463,966	148,187
California	5,111,205	1,545,300
Colorado	552,126	143,486
Connecticut	326,960	89,324
Delaware	116,181	32,862
District of Columbia	104,105	29,285
Florida	2,866,804	829,310
Georgia	1,503,877	513,320
Hawaii	126,481	32,442
Idaho	203,031	60,421
Illinois	2,543,257	470,646
Indiana	841,903	263,205
Iowa	319,716	85,662
Kansas	325,420	97,150

QUERIES

- 1. How many people crossed the US border between the year 2018-2019?**
- 2. How many people are below the poverty line in every state of the US?**
- 3. How many countries lie under high income group?**
- 4. Which item type product was highly purchased over the time in 2018 by each state of US?**
- 5. Compare the number of containers crossing US state border from last 5 years to today.**
- 6. Graph visualization of immigration vs population.**
- 7. Name the top 5 richest states in the US.**

(Horizontal view of datasets is attached below)

Horizontal View of Datasets

1. Income Table

1	Country Code	Region	IncomeGroup	Country
2	ABW	Latin America & Caribbean	High income	Aruba
3	AFG	South Asia	Low income	Afghanistan
4	AGO	Sub-Saharan Africa	Upper middle income	Angola
5	ALB	Europe & Central Asia	Upper middle income	Albania
6	AND	Europe & Central Asia	High income	Andorra
7	ARE	Middle East & North Africa	High income	United Arab Emirates
8	ARG	Latin America & Caribbean	Upper middle income	Argentina
9	ARM	Europe & Central Asia	Lower middle income	Armenia
10	ASM	East Asia & Pacific	Upper middle income	American Samoa
11	ATG	Latin America & Caribbean	High income	Antigua and Barbuda
12	AUS	East Asia & Pacific	High income	Australia
13	AUT	Europe & Central Asia	High income	Austria
14	AZE	Europe & Central Asia	Upper middle income	Azerbaijan
15	BDI	Sub-Saharan Africa	Low income	Burundi
16	BEL	Europe & Central Asia	High income	Belgium
17	BEN	Sub-Saharan Africa	Low income	Benin
18	BFA	Sub-Saharan Africa	Low income	Burkina Faso
19	BGD	South Asia	Lower middle income	Bangladesh
20	BGR	Europe & Central Asia	Upper middle income	Bulgaria

3. Immigration Table

1	Port Name	State	Port Code	Border	Date	Container	Value	Location
2	Calexico East	California	2507	US-Mexico Border	3/2/19 0:00	Trucks	34447	POINT (-115.48433000000001 32.67524)
3	Van Buren	Maine	108	US-Canada Border	3/1/19 0:00	Rail Containers Full	428	POINT (-67.94271 47.16207)
4	Otay Mesa	California	2506	US-Mexico Border	3/1/19 0:00	Trucks	81217	POINT (-117.05333 32.57333)
5	Nogales	Arizona	2604	US-Mexico Border	3/1/19 0:00	Trains	62	POINT (-110.93361 31.340279999999996)
6	Trout River	New York	715	US-Canada Border	3/2/19 0:00	Personal Vehicle Passengers	16377	POINT (-73.44253 44.990100000000005)
7	Madawaska	Maine	109	US-Canada Border	3/1/19 0:00	Trucks	179	POINT (-68.3271 47.35446)
8	Pembina	North Dakota	3401	US-Canada Border	3/1/19 0:00	Bus Passengers	1054	POINT (-97.24333 48.96639)
9	Progreso	Texas	2309	US-Mexico Border	3/1/19 0:00	Truck Containers Empty	1808	POINT (-97.94889 26.061670000000003)
10	Portal	North Dakota	3403	US-Canada Border	3/2/19 0:00	Rail Containers Empty	6685	POINT (-102.54917 48.99583)
11	Champlain-Rouses Point	New York	712	US-Canada Border	3/1/19 0:00	Trucks	24759	POINT (-73.44694 44.98639)
12	Opheim	Montana	3317	US-Canada Border	3/1/19 0:00	Personal Vehicles	235	POINT (-106.40265 48.85574)
13	Neché	North Dakota	3404	US-Canada Border	3/1/19 0:00	Trucks	584	POINT (-97.55444 48.98611)
14	Lancaster	Minnesota	3430	US-Canada Border	3/1/19 0:00	Personal Vehicles	1612	POINT (-96.82024 48.849110000000001)
15	Derby Line	Vermont	209	US-Canada Border	3/1/19 0:00	Buses	80	POINT (-72.09944 45.005)
16	Sarles	North Dakota	3409	US-Canada Border	3/1/19 0:00	Personal Vehicle Passengers	509	POINT (-98.99457 48.941050000000004)
17	Wildhorse	Montana	3323	US-Canada Border	3/1/19 0:00	Truck Containers Full	122	POINT (-109.67761 48.54863)
18	Derby Line	Vermont	209	US-Canada Border	3/1/19 0:00	Truck Containers Full	6483	POINT (-72.09944 45.005)
19	Lynden	Washington	3023	US-Canada Border	3/1/19 0:00	Pedestrians	79	POINT (-122.44316000000002 48.94802)
20	Vanceboro	Maine	105	US-Canada Border	3/1/19 0:00	Pedestrians	3	POINT (-67.42955 45.55984)

2. Product Table

1	Region	Country	Item Type
2	Middle East and North Africa	Libya	Cosmetics
3	North America	Canada	Vegetables
4	Middle East and North Africa	Libya	Baby Food
5	Asia	Japan	Cereal
6	Sub-Saharan Africa	Chad	Fruits
7	Europe	Armenia	Cereal
8	Sub-Saharan Africa	Eritrea	Cereal
9	Europe	Montenegro	Clothes
10	Central America and the Caribbean	Jamaica	Vegetables
11	Australia and Oceania	Fiji	Vegetables
12	Sub-Saharan Africa	Togo	Clothes
13	Europe	Montenegro	Snacks
14	Europe	Greece	Household
15	Sub-Saharan Africa	Sudan	Cosmetics
16	Asia	Maldives	Fruits
17	Europe	Montenegro	Clothes
18	Europe	Estonia	Office Supplies
19	North America	Greenland	Beverages
20	Sub-Saharan Africa	Cape Verde	Clothes

4. Population Table

COUNTRY USA	
State	2018 Population
California	39776830
Texas	28704330
Florida	21312211
New York	19862512
Pennsylvania	12823989
Illinois	12768320
Ohio	11694664
Georgia	10545138
North Carolina	10390149
Michigan	9991177
New Jersey	9032872
Virginia	8525660
Washington	7530552
Arizona	7123898
Massachusetts	6895917
Tennessee	6782564
Indiana	6699629
Missouri	6135888

5. Poverty Table

USA	People in Poverty		Children in Poverty	
	Lower Level	Upper Level	Lower Bound	Upper Bound
Alabama	784,517	820,009	253,694	272,124
Alaska	75,273	83,583	23,957	28,201
Arizona	1,002,773	1,043,901	324,723	347,577
Arkansas	463,966	487,800	149,187	162,569
California	5,111,205	5,217,133	1,586,526	1,645,300
Colorado	552,126	580,110	143,486	160,312
Connecticut	326,360	350,942	89,324	101,814
Delaware	116,181	127,003	32,862	37,720
District of Columbia	104,105	114,899	29,285	35,421
Florida	2,860,804	2,942,948	828,310	872,538
Georgia	1,503,877	1,568,261	513,320	550,706
Hawaii	126,481	137,921	32,442	38,702
Idaho	203,631	220,243	60,371	69,421
Illinois	1,543,257	1,594,083	470,646	500,886
Indiana	841,903	879,603	263,205	285,029
Iowa	319,716	337,542	85,662	94,830
Kansas	325,420	345,322	97,150	108,566