

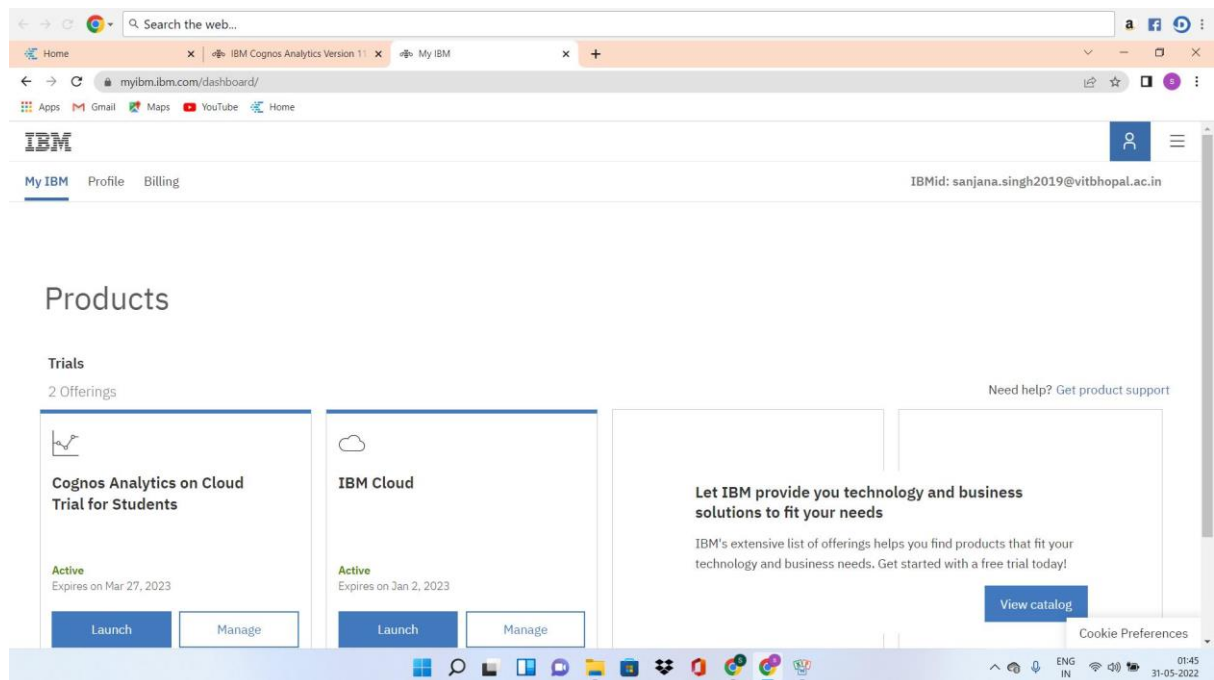
Name: Sanjana Singh

Registration No.: 19BCE10274

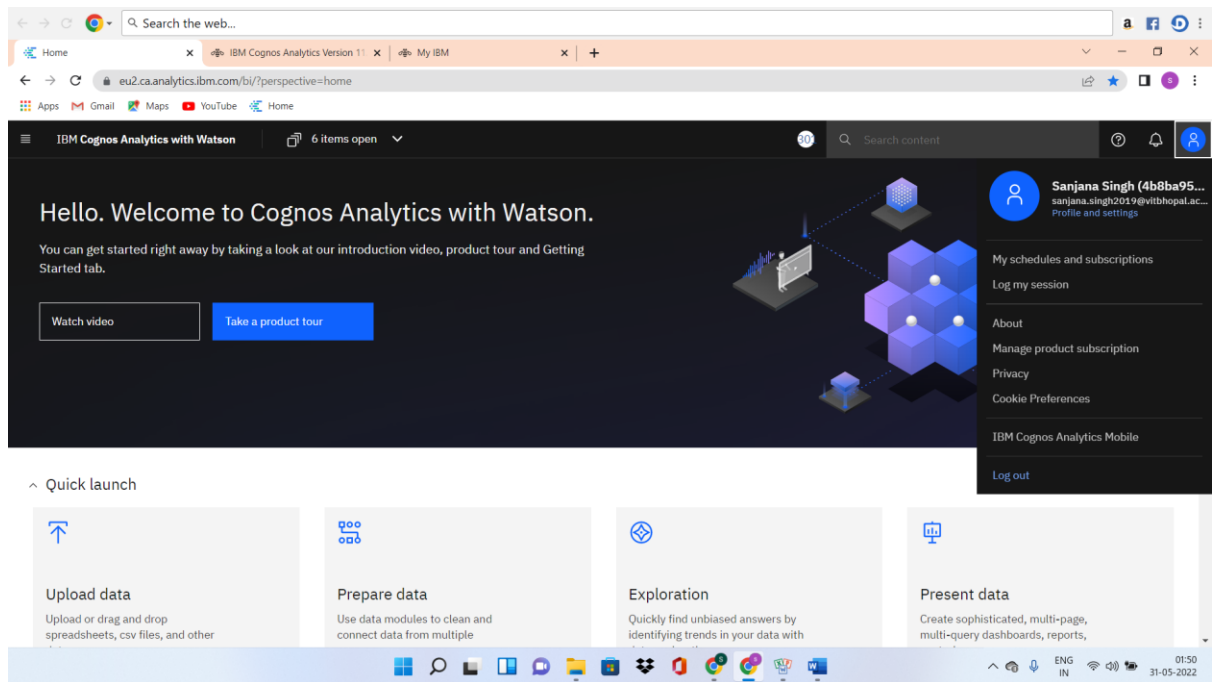
Mentor's Name: Uma Rani

World Population Data Analytics using IBM Cognos

1. IBM Cloud Creation



2. IBM Cognos Analytics

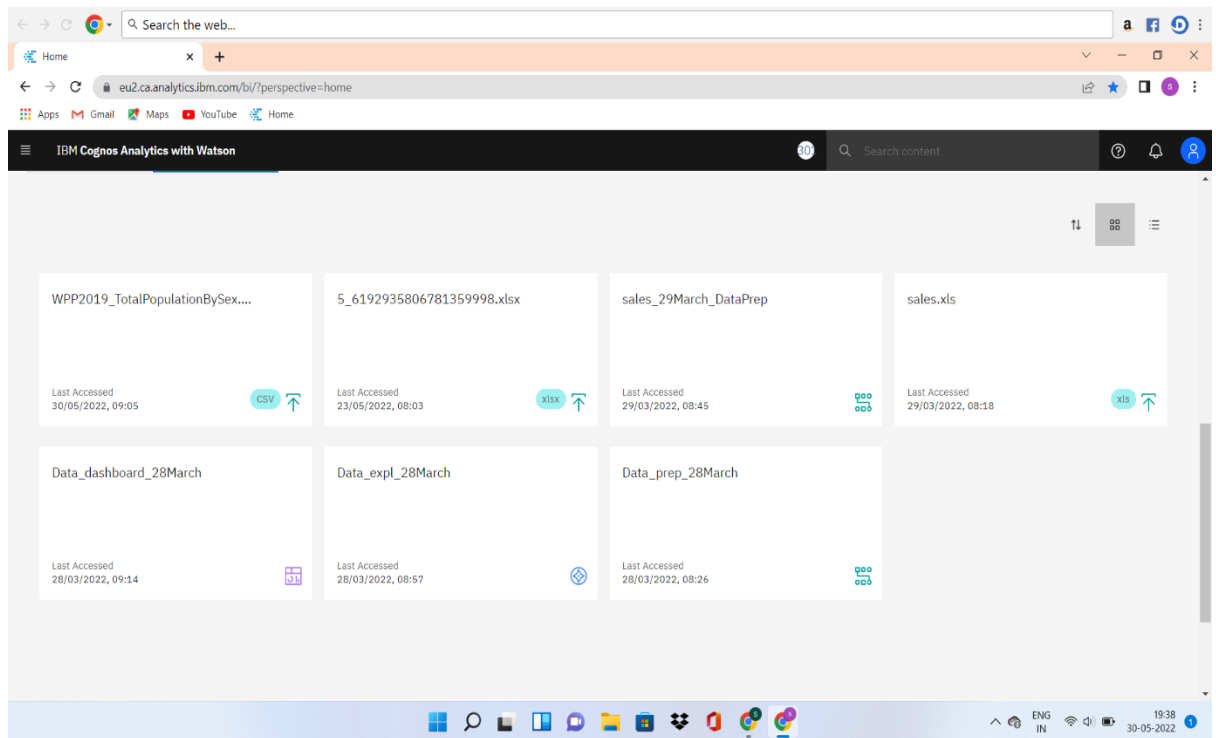


3. Working with the dataset

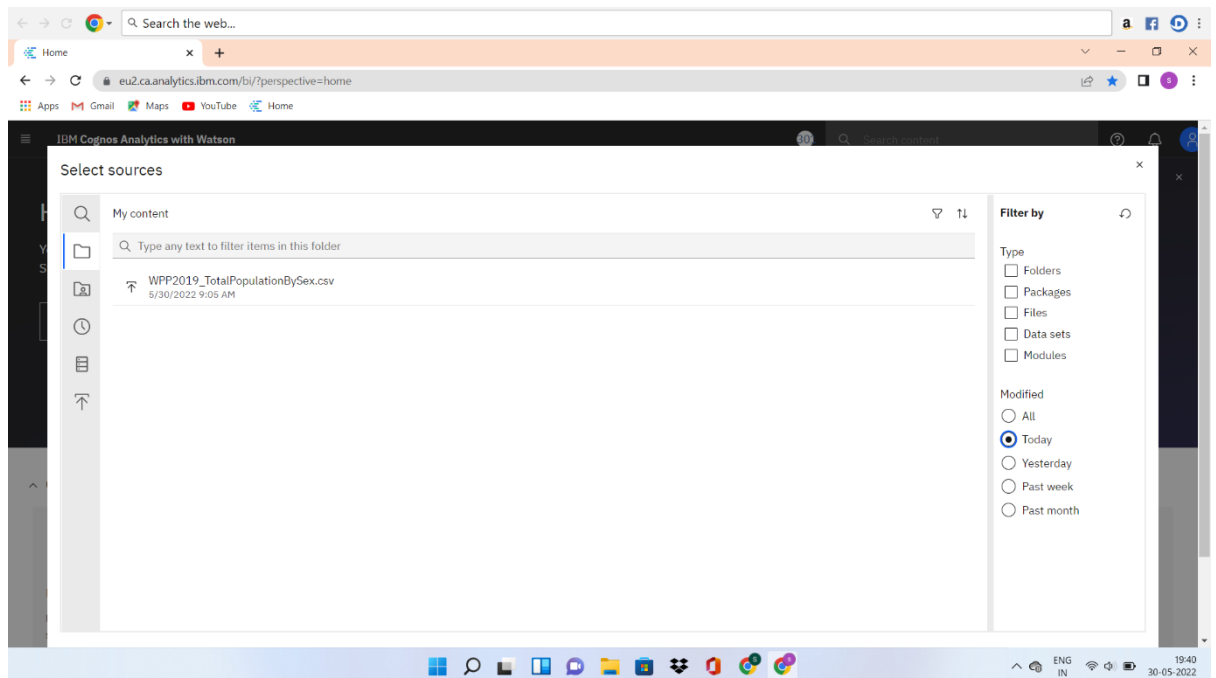
3.1 Understand the dataset

WPP2019_TotalPopulationBySex - Microsoft Excel (Product Activation Failed)																		
File Home Insert Page Layout Formulas Data Review View																		
Clipboard Font Alignment Number Styles Cells Editing																		
A1 LocID																		
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	LocID	Location	VarID	Variant	Time	MidPeriod	PopMale	PopFemal	PopTotal	PopDensity								
2	4	Afghanistan	2	Medium	1950	1950.5	4099.243	3652.874	7752.117	11.874								
3	4	Afghanistan	2	Medium	1951	1951.5	4134.756	3705.395	7840.151	12.009								
4	4	Afghanistan	2	Medium	1952	1952.5	4174.45	3761.546	7935.996	12.156								
5	4	Afghanistan	2	Medium	1953	1953.5	4218.336	3821.348	8039.684	12.315								
6	4	Afghanistan	2	Medium	1954	1954.5	4266.484	3884.832	8151.316	12.486								
7	4	Afghanistan	2	Medium	1955	1955.5	4318.945	3952.047	8270.992	12.669								
8	4	Afghanistan	2	Medium	1956	1956.5	4375.8	4023.073	8398.873	12.865								
9	4	Afghanistan	2	Medium	1957	1957.5	4437.157	4098	8535.157	13.073								
10	4	Afghanistan	2	Medium	1958	1958.5	4503.156	4176.941	8680.097	13.295								
11	4	Afghanistan	2	Medium	1959	1959.5	4573.914	4260.033	8833.947	13.531								
12	4	Afghanistan	2	Medium	1960	1960.5	4649.573	4347.394	8996.967	13.781								
13	4	Afghanistan	2	Medium	1961	1961.5	4730.25	4439.156	9169.406	14.045								
14	4	Afghanistan	2	Medium	1962	1962.5	4816.05	4535.392	9351.442	14.324								
15	4	Afghanistan	2	Medium	1963	1963.5	4907.03	4636.17	9543.2	14.618								
16	4	Afghanistan	2	Medium	1964	1964.5	5003.245	4741.527	9744.772	14.926								
17	4	Afghanistan	2	Medium	1965	1965.5	5104.765	4851.553	9956.318	15.25								
18	4	Afghanistan	2	Medium	1966	1966.5	5210.122	4964.718	10174.84	15.585								
19	4	Afghanistan	2	Medium	1967	1967.5	5319.123	5080.813	10399.94	15.93								
20	4	Afghanistan	2	Medium	1968	1968.5	5434.458	5202.606	10637.06	16.293								
21	4	Afghanistan	2	Medium	1969	1969.5	5559.836	5333.936	10893.77	16.686								
22	4	Afghanistan	2	Medium	1970	1970.5	5697.024	5476.63	11173.65	17.115								
23	4	Afghanistan	2	Medium	1971	1971.5	5845.351	5630.099	11475.45	17.577								
24	4	Afghanistan	2	Medium	1972	1972.5	6000.805	5760.277	11761.08	18.061								

3.2 Loading the dataset



3.3 Prepare the datasets



IBM Cognos Analytics with Watson

Search content

Properties

Data module

Search

New data module

- Navigation paths
- WPP2019_T...BySex.csv
 - Row Id
 - LocID
 - Location
 - VarID
 - Variant
 - Time
 - MidPeriod
 - PopMale
 - PopFemale
 - PopTotal
 - PopDensity

Grid Relationships Custom tables

Preview data

To preview data, select a table, a column in a table, or a folder that contains columns.

IBM Cognos Analytics with Watson

Search content

Properties

Data module

Search

New data module

- Navigation paths
- WPP2019_T...BySex.csv
 - Row Id
 - LocID
 - Location
 - VarID
 - Variant
 - Time
 - MidPeriod
 - PopMale
 - PopFemale
 - PopTotal
 - PopDensity

Grid Relationships Custom tables

Row Id	LocID	Location	VarID	Variant	Time	MidPeriod
1	4	Afghanistan	2	Medium	1950	1950.5
2	4	Afghanistan	2	Medium	1951	1951.5
3	4	Afghanistan	2	Medium	1952	1952.5
4	4	Afghanistan	2	Medium	1953	1953.5
5	4	Afghanistan	2	Medium	1954	1954.5
6	4	Afghanistan	2	Medium	1955	1955.5
7	4	Afghanistan	2	Medium	1956	1956.5
8	4	Afghanistan	2	Medium	1957	1957.5
9	4	Afghanistan	2	Medium	1958	1958.5
10	4	Afghanistan	2	Medium	1959	1959.5
11	4	Afghanistan	2	Medium	1960	1960.5
12	4	Afghanistan	2	Medium	1961	1961.5

- Removing the existing PopTotal column

The screenshot shows the IBM Cognos Analytics interface. On the left, the 'Data module' pane lists various fields including 'PopTotal'. The main area displays a table with columns: Variant, Time, MidPeriod, PopMale, PopFemale, PopTotal, and PopDensity. A context menu is open over the 'PopTotal' column, showing options like 'Filter...', 'Create calculation...', 'Create data group...', 'Hide from users', 'Remove', 'Refresh properties...', 'Format data...', 'Clean...', 'Rename', 'Cut', 'Copy', and 'Properties'.

Variant	Time	MidPeriod	PopMale	PopFemale	PopTotal	PopDensity
1955	1955.5	4318.945	3952.047	8270.992	12.669	
1956	1956.5	4375.8	4023.073	8398.873	12.865	
1957	1957.5	4437.157	4098	8535.157	13.073	
1958	1958.5	4503.156	4176.941	8680.097	13.295	
1959	1959.5	4573.914	4260.033	8833.947	13.531	
1960	1960.5	4649.573	4347.394	8996.967	13.781	
1961	1961.5	4730.25	4439.156	9169.406	14.045	
1962	1962.5	4816.05	4535.392	9351.442	14.324	
1963	1963.5	4907.03	4636.17	9543.2	14.618	
1964	1964.5	5003.245	4741.527	9744.772	14.926	
1965	1965.5	5104.765	4851.553	9956.318	15.25	
1966	1966.5	5210.122	4964.718	10174.84	15.585	

- Calculating the new PopTotal column. $\text{PopTotal} = \text{PopMale} + \text{PopFemale}$ and validate the expression to formate the data.

The screenshot shows the 'Create calculation' dialog in IBM Cognos Analytics. The 'Name' field is 'C_PopTotal_New'. The 'Expression' field contains 'PopMale + PopFemale'. The 'Validation Results' section shows a green checkmark and the message 'The expression is valid.'

- Creation of new PopTotal column

The screenshot shows the IBM Cognos Analytics web interface. On the left, the 'Data module' pane lists 'C_PopTotal_New' under 'WPP2019_T...BySex.csv'. The main area displays a data grid with the following columns: C_PopTotal_New, Row Id, LocID, Location, VarID, Variant, and Time. The data represents population totals for Afghanistan from 1950 to 1961.

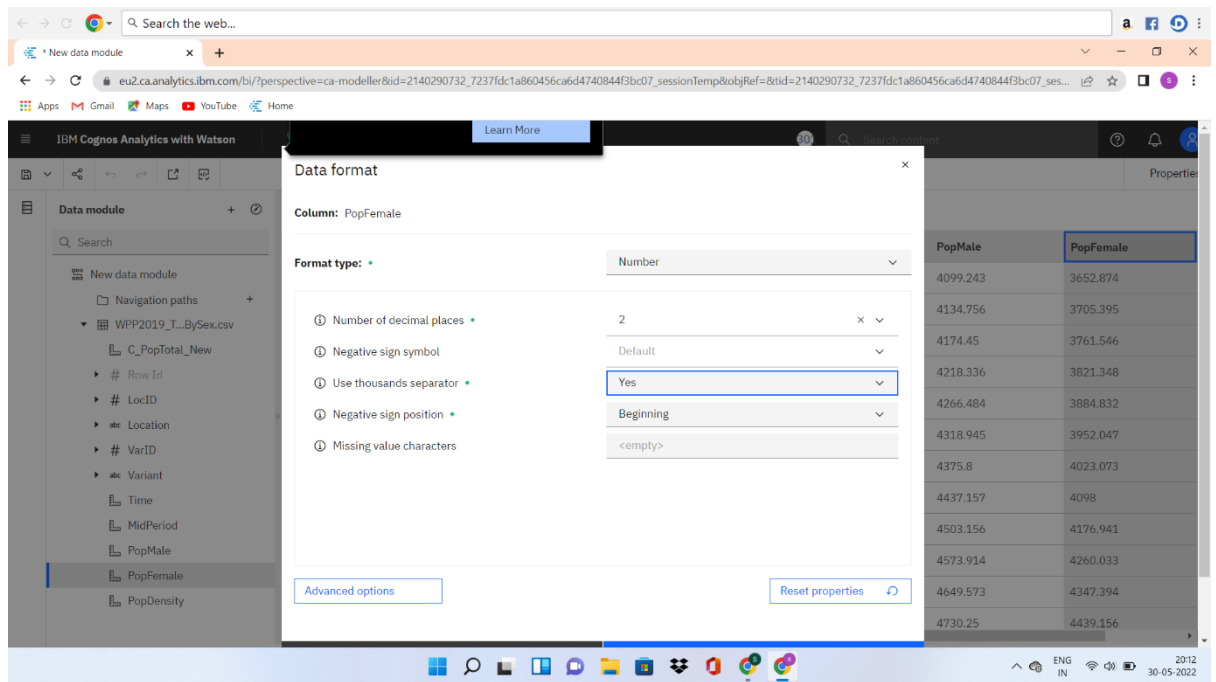
C_PopTotal_New	Row Id	LocID	Location	VarID	Variant	Time
7752.117	1	4	Afghanistan	2	Medium	1950
7840.151	2	4	Afghanistan	2	Medium	1951
7935.995999999999	3	4	Afghanistan	2	Medium	1952
8039.684	4	4	Afghanistan	2	Medium	1953
8151.316000000001	5	4	Afghanistan	2	Medium	1954
8270.992	6	4	Afghanistan	2	Medium	1955
8398.873	7	4	Afghanistan	2	Medium	1956
8535.157	8	4	Afghanistan	2	Medium	1957
8680.097	9	4	Afghanistan	2	Medium	1958
8833.947	10	4	Afghanistan	2	Medium	1959
8996.967	11	4	Afghanistan	2	Medium	1960
9169.405999999999	12	4	Afghanistan	2	Medium	1961

- Formate the PopMale data type

The screenshot shows the 'Data format' dialog box for the 'PopMale' column. The 'Format type' is set to 'Number'. The 'Number of decimal places' is set to 2. The 'Negative sign symbol' is 'Default', 'Use thousands separator' is 'Yes', 'Negative sign position' is 'Beginning', and 'Missing value characters' are '<empty>'. In the background, a data table shows 'MidPeriod' and 'PopMale' values.

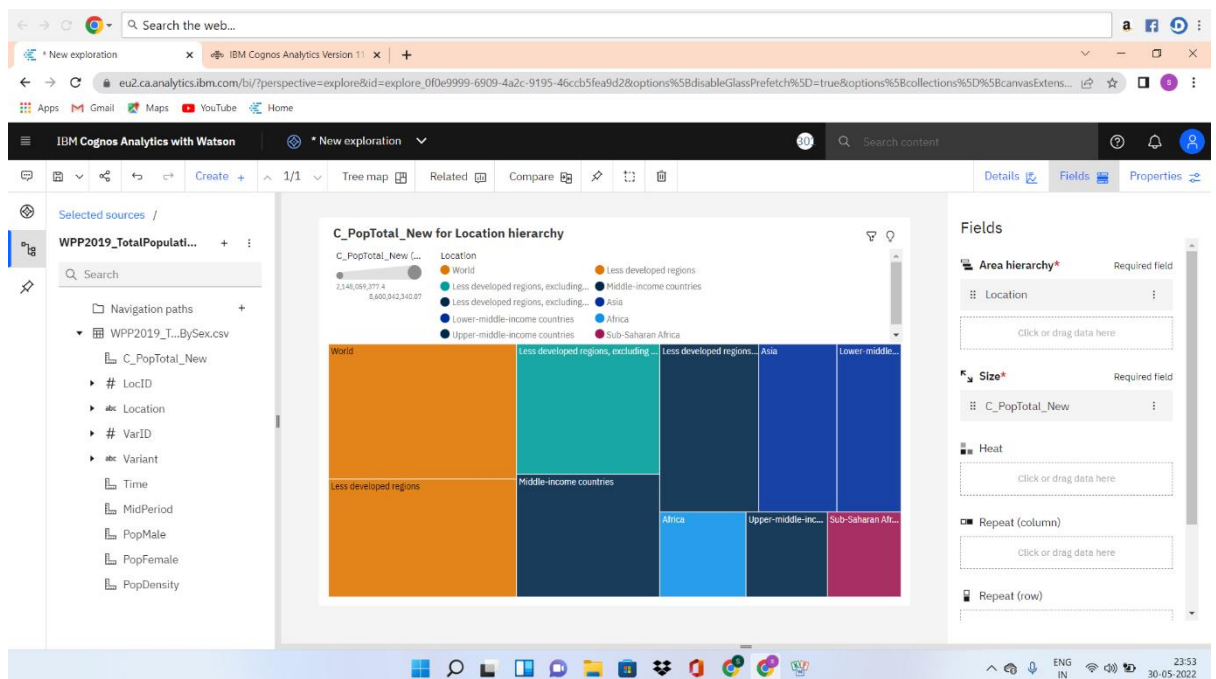
MidPeriod	PopMale
1950.5	4099.243
1951.5	4134.756
1952.5	4174.45
1953.5	4218.336
1954.5	4266.484
1955.5	4318.945
1956.5	4375.8
1957.5	4437.157
1958.5	4503.156
1959.5	4573.914
1960.5	4649.573
1961.5	4730.25

- Formate the PopFemale data type

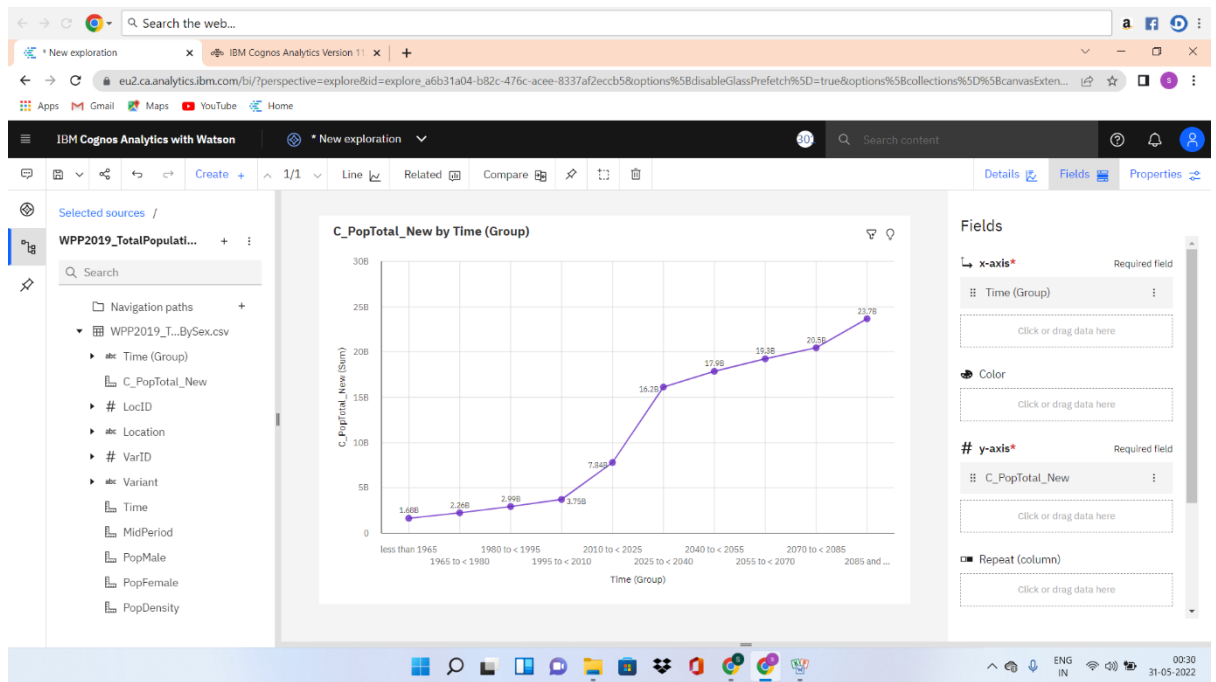


2. Data Visualization Chart

4.1 Top10 Pop Total by Location Using Tree Map

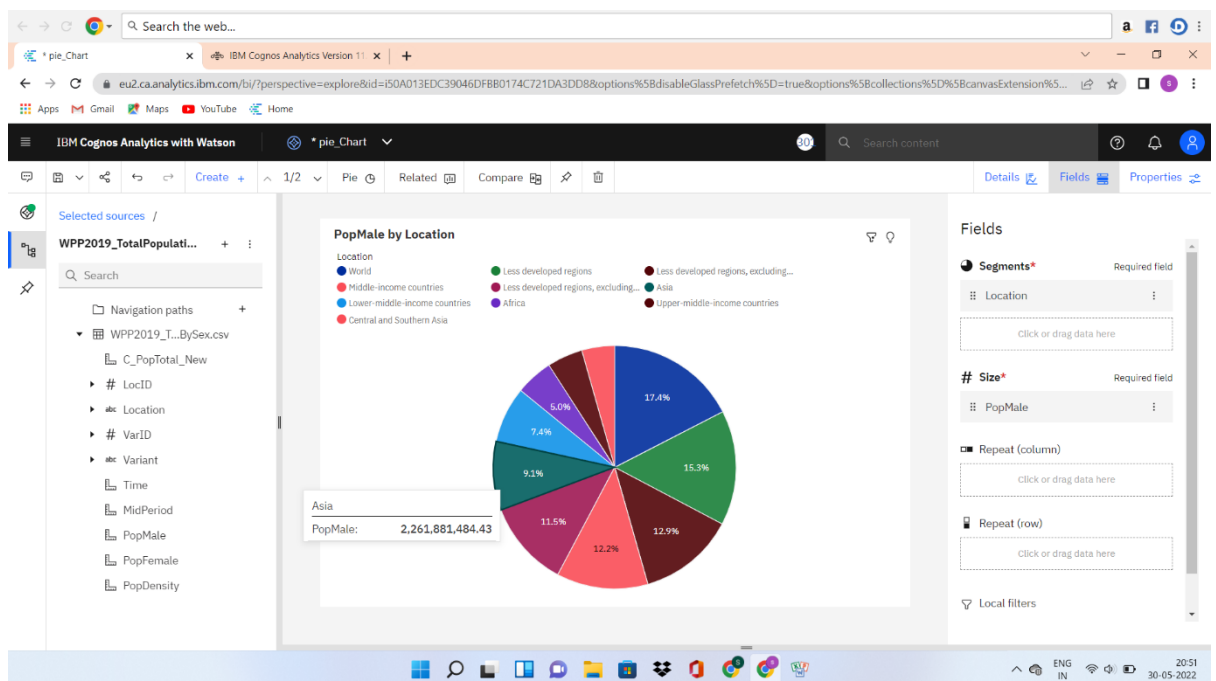


4.2 Pop Total by Time Using Line Chart

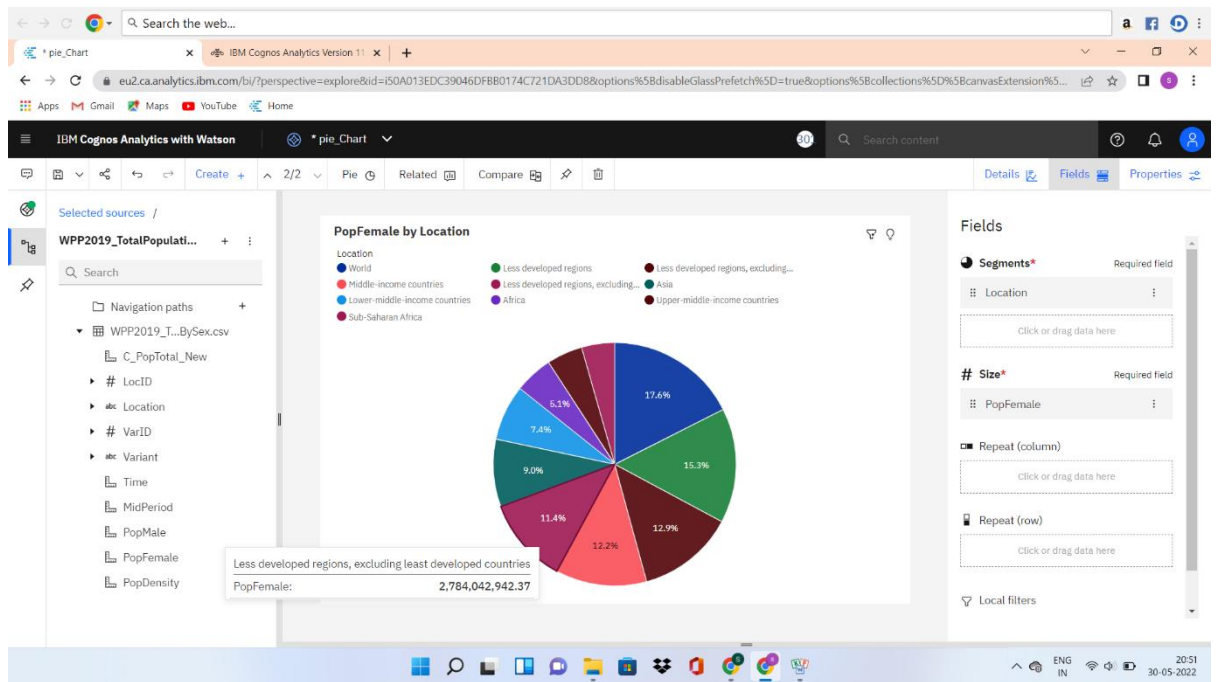


4.3 Pop Male by Location and Pop Female by Location using Pie Charts

- PopMale by Location

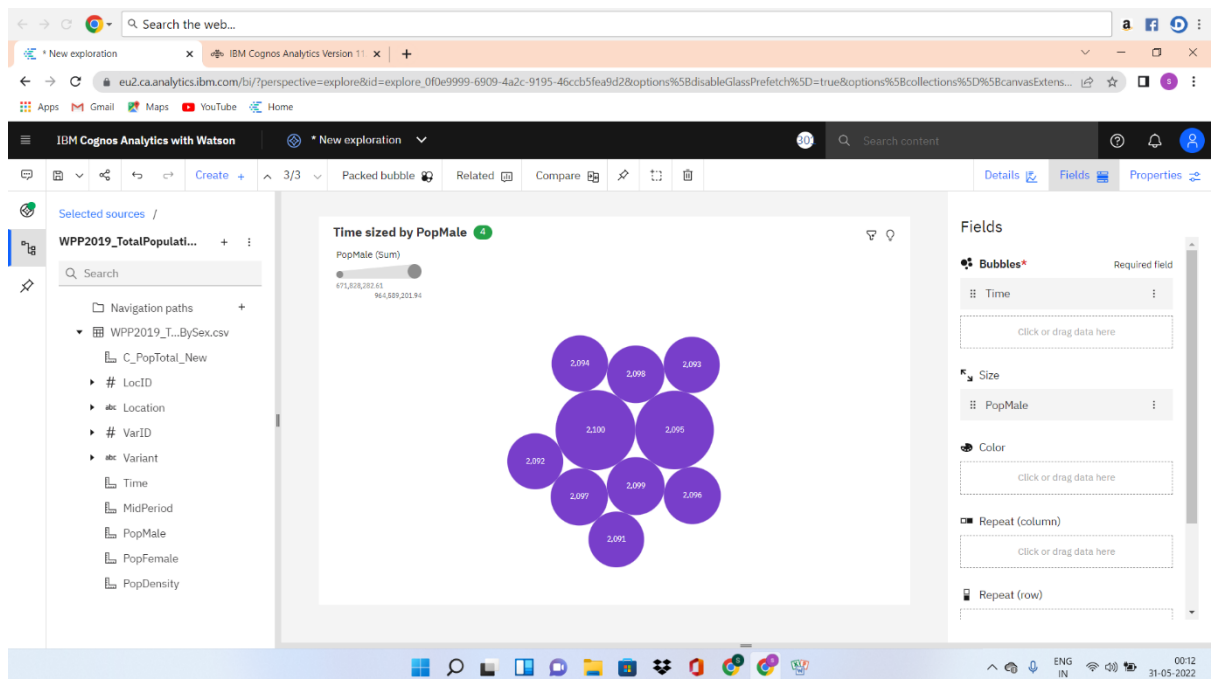


- PopFemale by Location

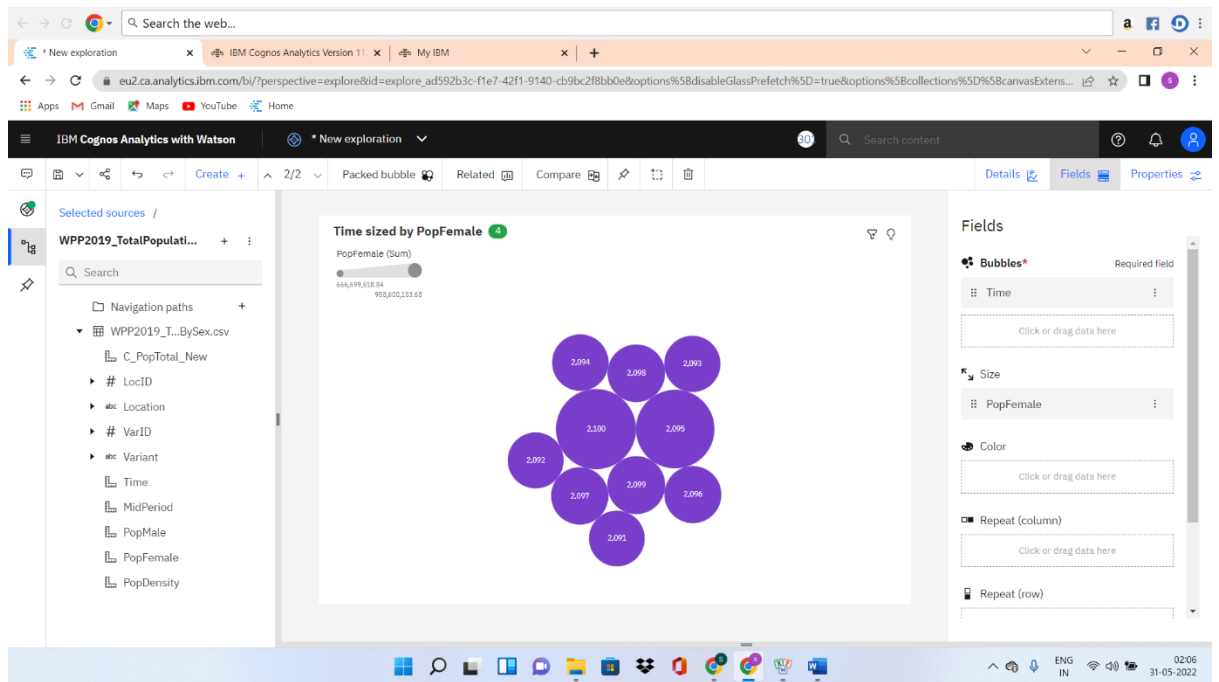


4.4 Pop Male by Time and Pop Female Using Packed Bubble Charts

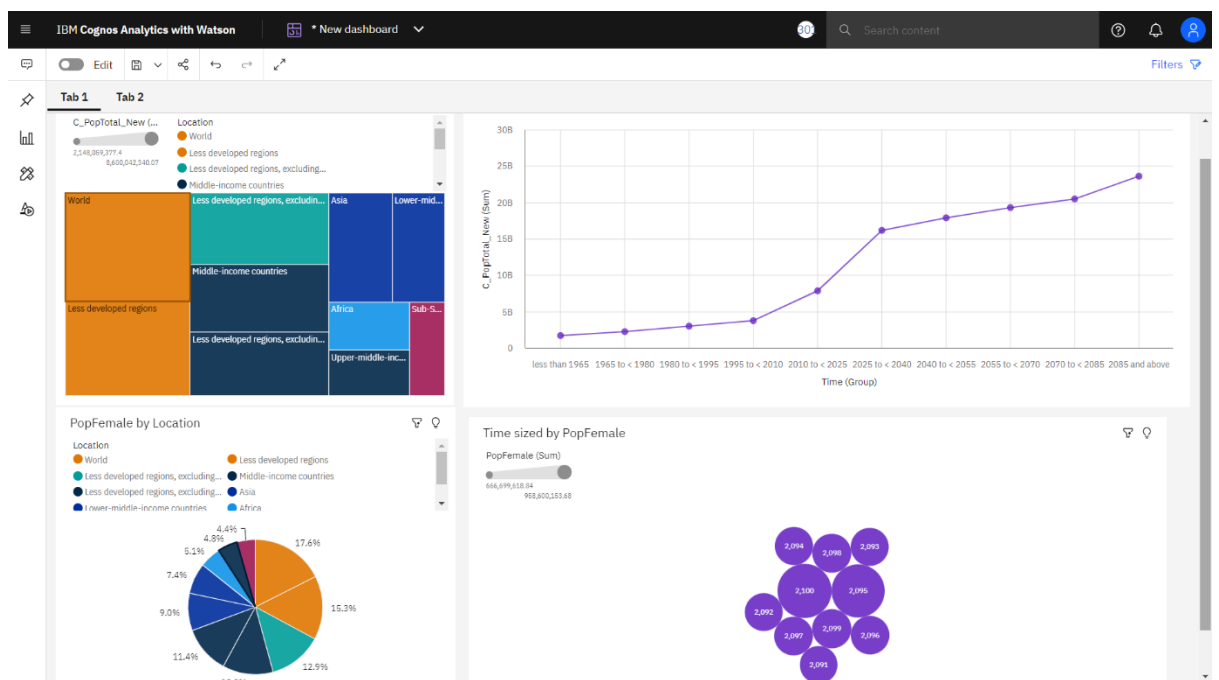
- Time sized by PopMale



- Time sized by PopFemale



4.5 Building of Dashboard



4.6 Pop Male, Pop Female and Pop Total using Summary

