**DATA WAREHOUSE**

**NAAN MUDHALVAN PROJECT ON**

**HEALTH­-CARE**

**USE CASE: Health Care**

**Problem Statements :**

* Find the people’s age less than 50 who are addicted to smoke.
* Find the humidity of pollution which is in specific period.
* Find the table of admission data where the column of 20 is null.
* Find the total count for single column in specific table.
* Find the Quality of air between specific range of period.

**Table Contains :**

* HDHI Pollution Table.
* HDHI Mortality Table.
* HDHI Admission Table.

**Creation of data warehouse, database,schema and table in snowflake:**

Create warehouse chendhuran\_college;

Create database chendhuran\_college\_cse;

Create schema cse\_A;

Create table HDHI Pollution (DATE varchar, AQI varchar, 'PM2.5 AVG' varchar, 'PM2.5 MIN'varchar, PM2.5 MAX varchar, 'PM10 AVG' varchar, 'PM10 MIN' varchar, 'PM10 MAX' varchar, 'NO2 AVG'varchar, 'NO2 MIN' varchar, 'NO2 MAX' varchar, 'NH3 AVG' varchar, 'NH3 MIN' varchar, 'NH3 MAX'varchar, 'SO2 AVG' varchar, 'SO2 MIN' varchar, 'SO2 MAX' varchar, 'CO AVG' varchar, 'CO MIN'varchar, 'CO MAX' varchar, 'OZONE AVG' varchar, 'OZONE MIN'varchar, 'OZONE MAX' varchar, 'PROMINENT POLLUTENT' varchar, 'MAX TEMP' varchar, 'MIN TEMP' varchar, 'HUMIDITY' varchar);

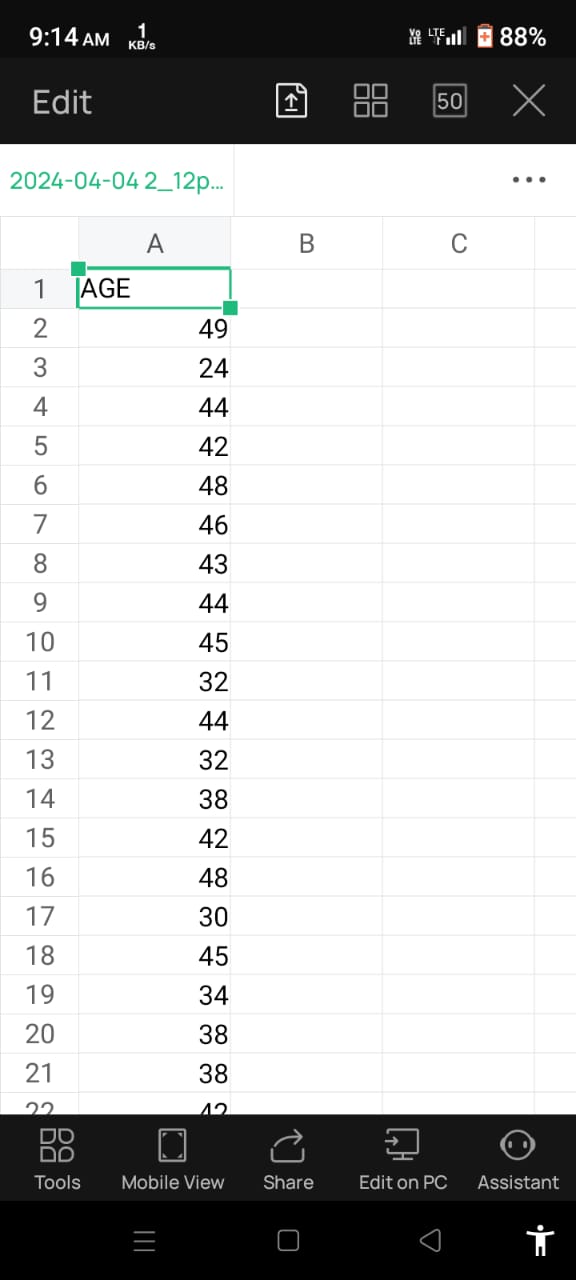
**Problem statement 1:**

Find the people’s age less than 50 who are addicted to smoke.

**Query:**

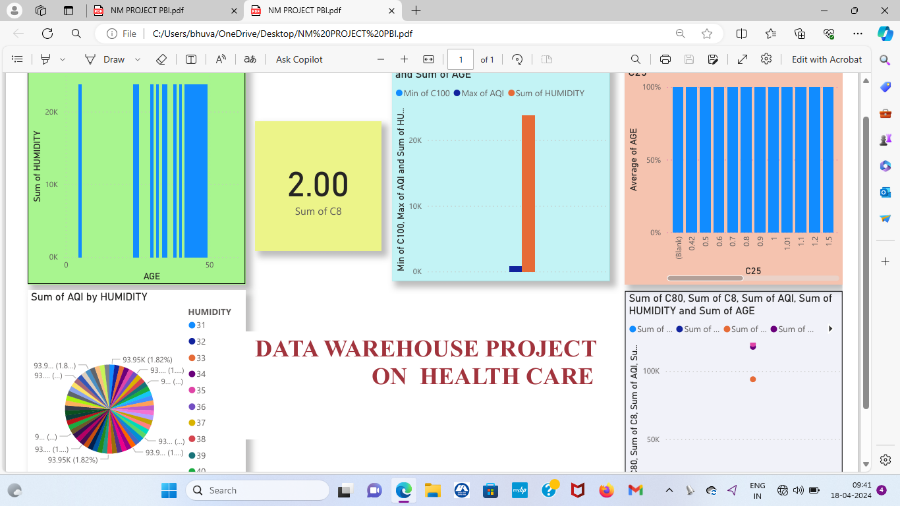
Select age from "HDHI Mortality Date" where age<50;

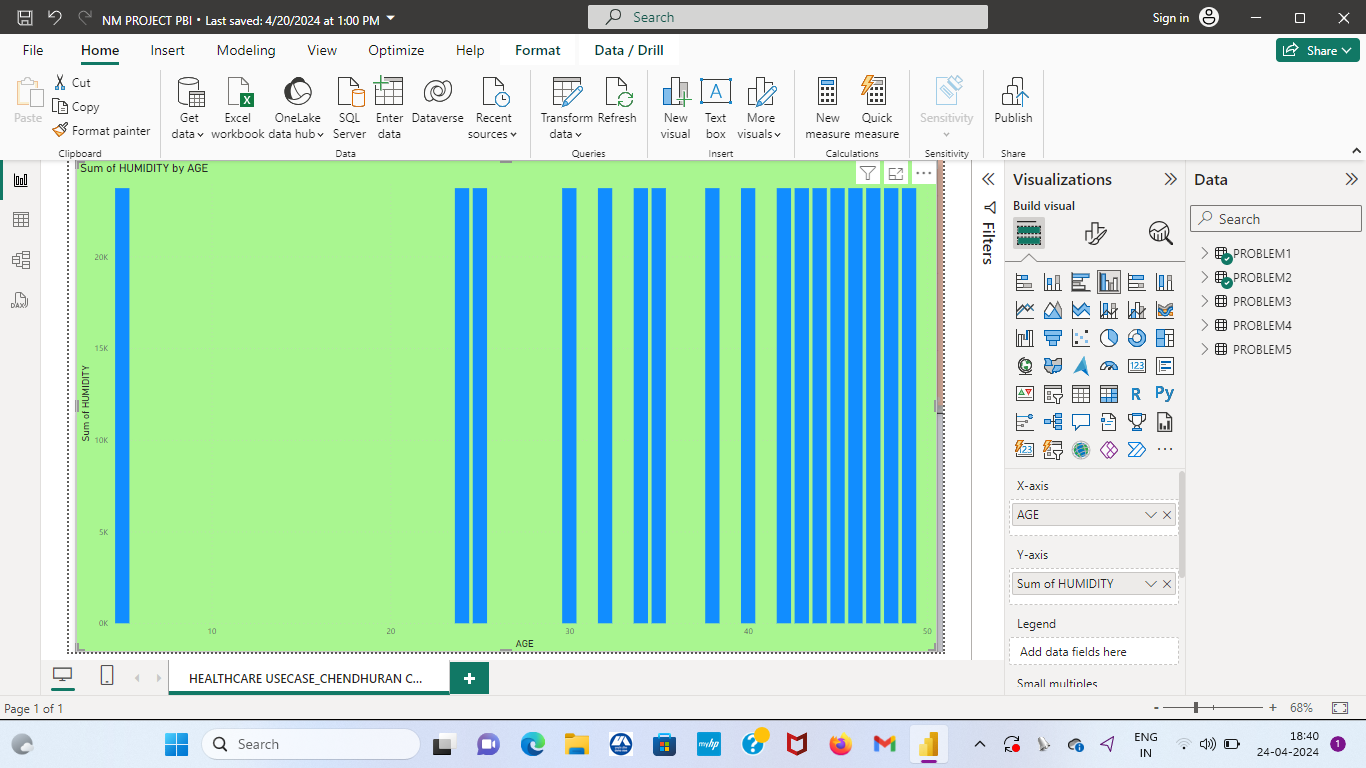
**Snowflake Output:**



**Power BI Output:**

Choose a report view – choose any one chart in visualizations – choose an age in the table – drag a column to the fields of x-axis and y-axis visualization—change the legends for colour change as your need.



****

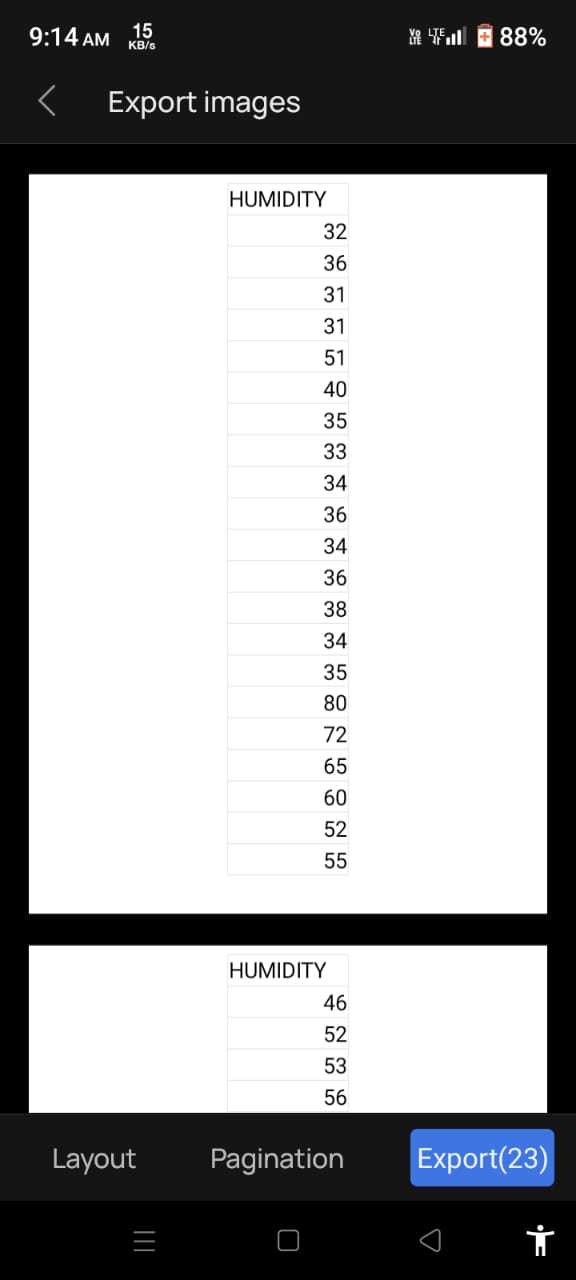
**Problem statement 2:**

Find the humidity of pollution which is in specific period.

**Query:**

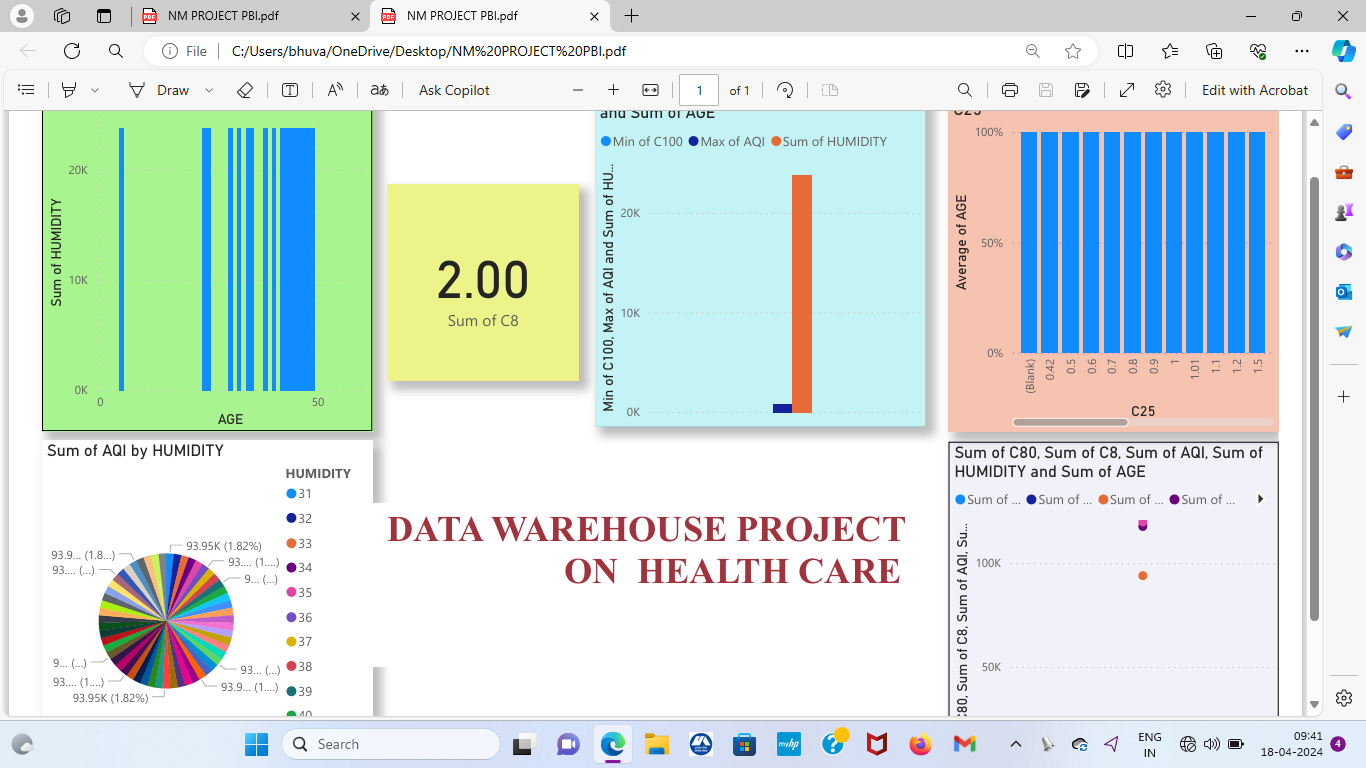
Select humidity from"HDHI Pollution Data"where humidity>30;

**Snowflake Output:**



**Power BI Output:**

Choose a report view – choose clustered column chart in visualizations – choose a humidity by sum of age in the table – drag a column to the fields of x-axis and y-axis visualization—change the legends for colour change as your need and also change the axis name of sum into avg,min,max etc..



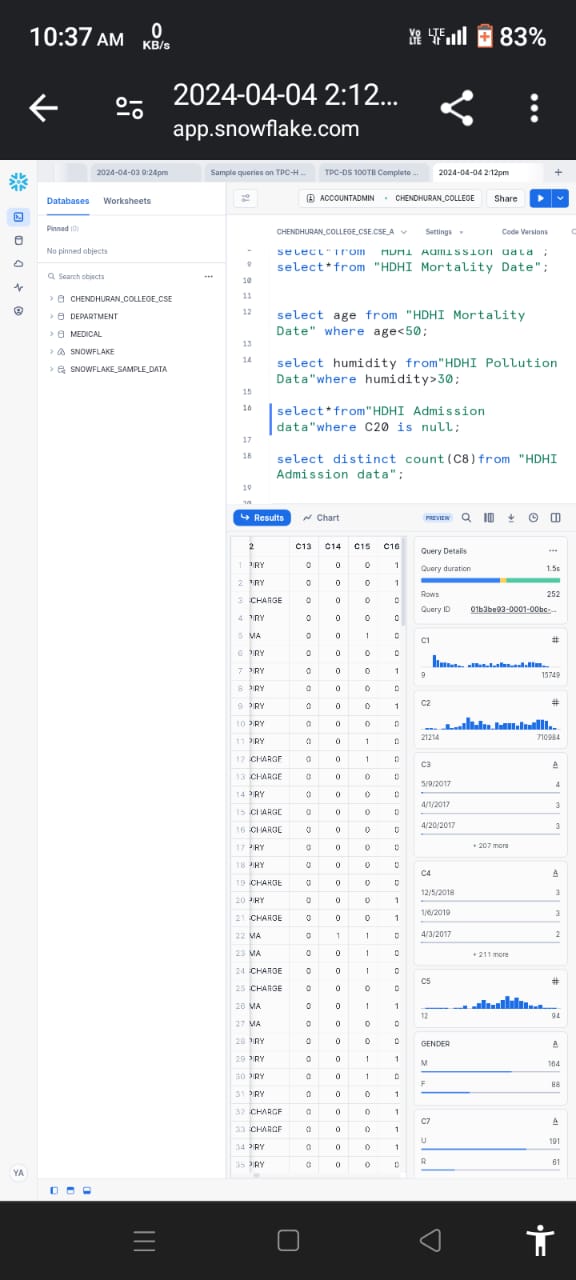
**Problem statement 3:**

Find the table of admission data where the column of 20 is null

**Query:**

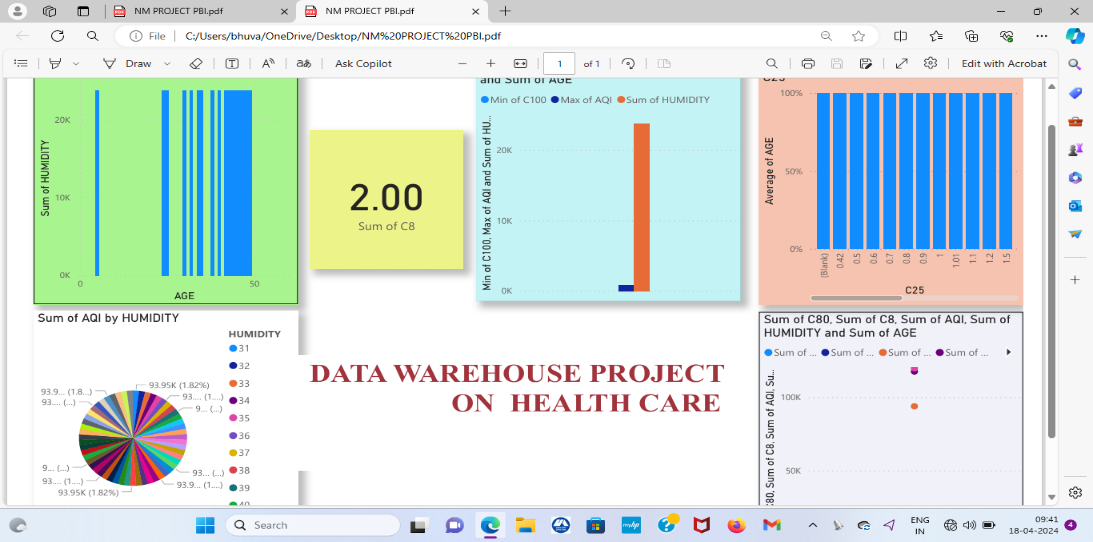
Select\*from"HDHI Admission data"where C20 is null;

**Snowflake Output:**

****

**Power BI output:**

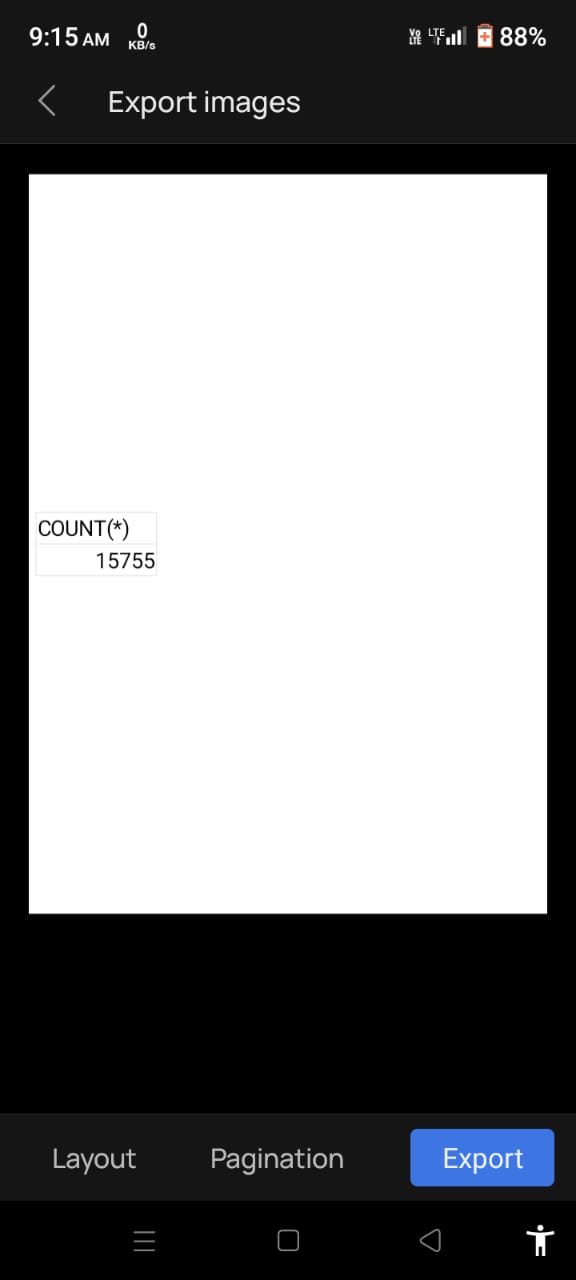
Choose a report view – choose any one chart in visualizations – choose a column of specific table which is in null condition – drag a column to the fields of x-axis and y-axis visualization—change the legends for colour, font change as your need.



**Problem statement 4:**

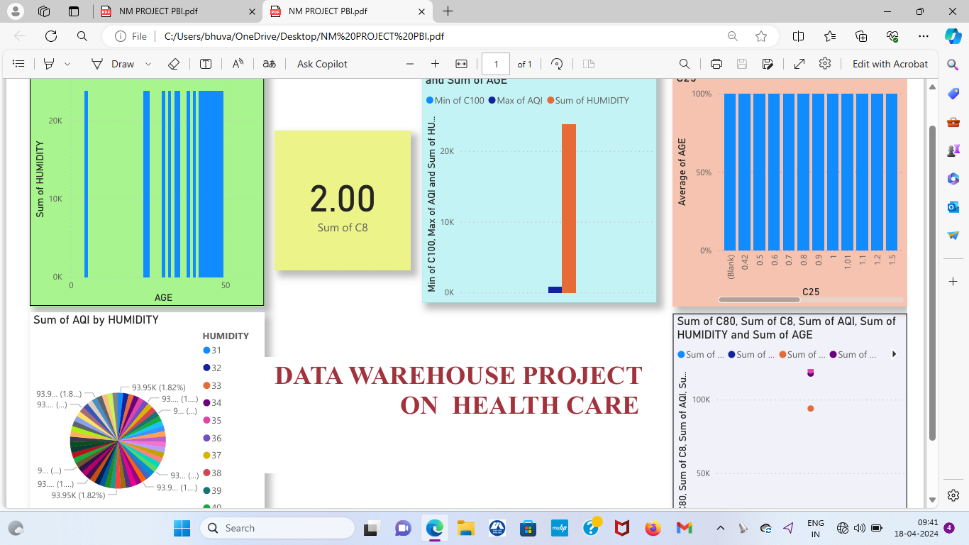
Find the total count for single column in specific table.

**Snowflake Output:**

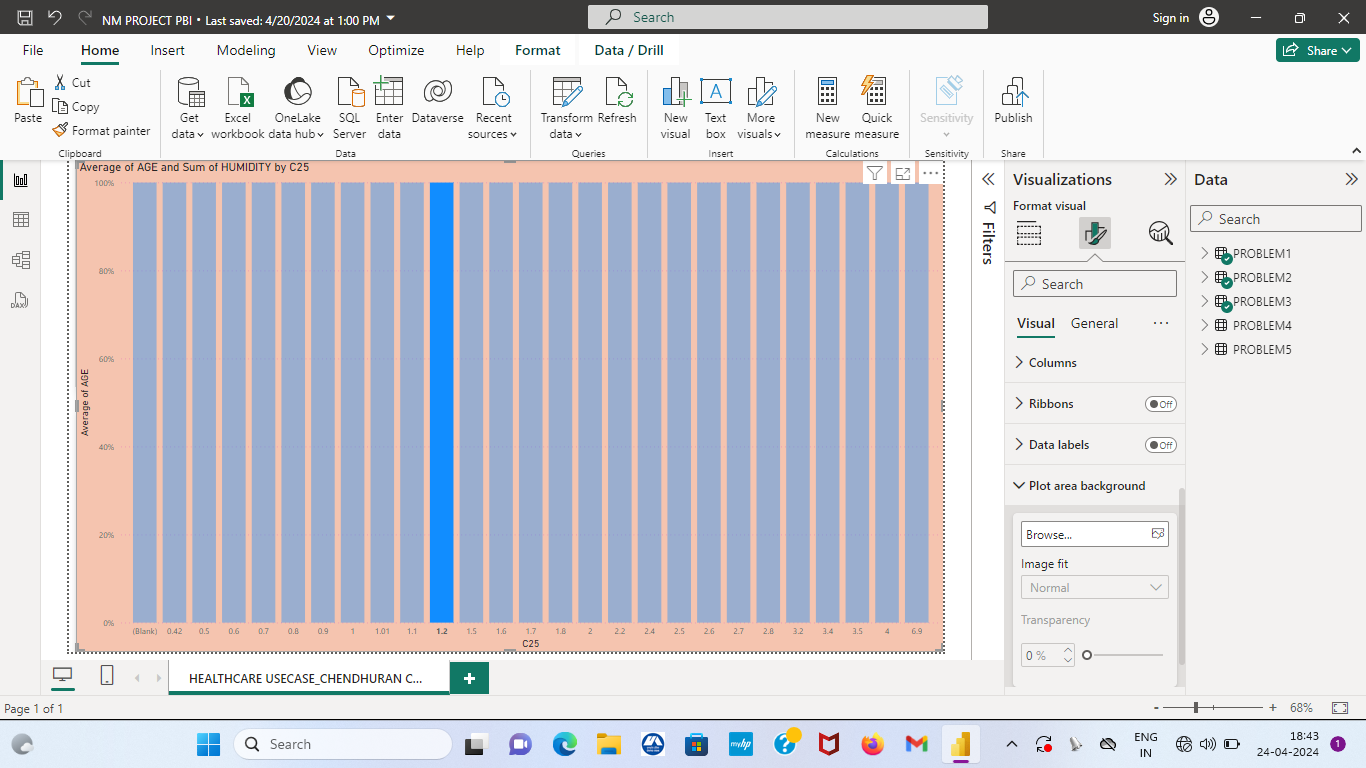


**Power BI output:**

Choose a report view – choose any one chart in visualizations – choose a count of specific column in the table – drag a column to the fields of x-axis and y-axis visualization—change the legends for colour change as your need.



**AVERAGE AND SUM BY SINGLE COLUMN OF C20**

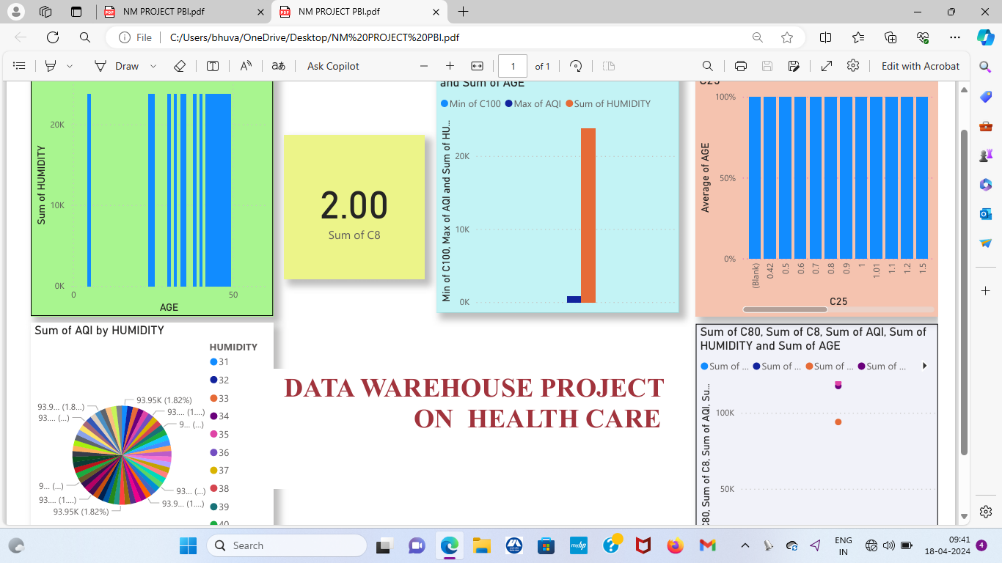


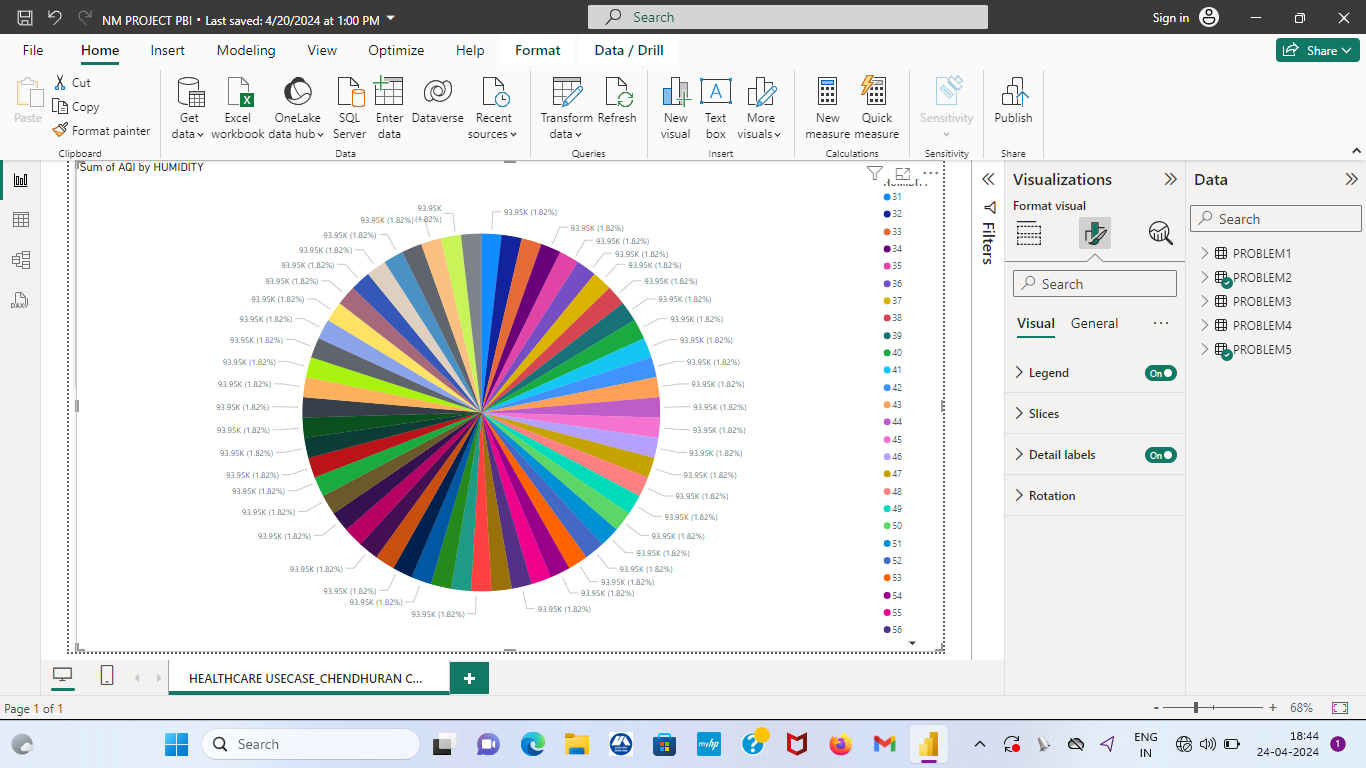
**Problem statement 5:**

Find the Quality of air between specific range of period.

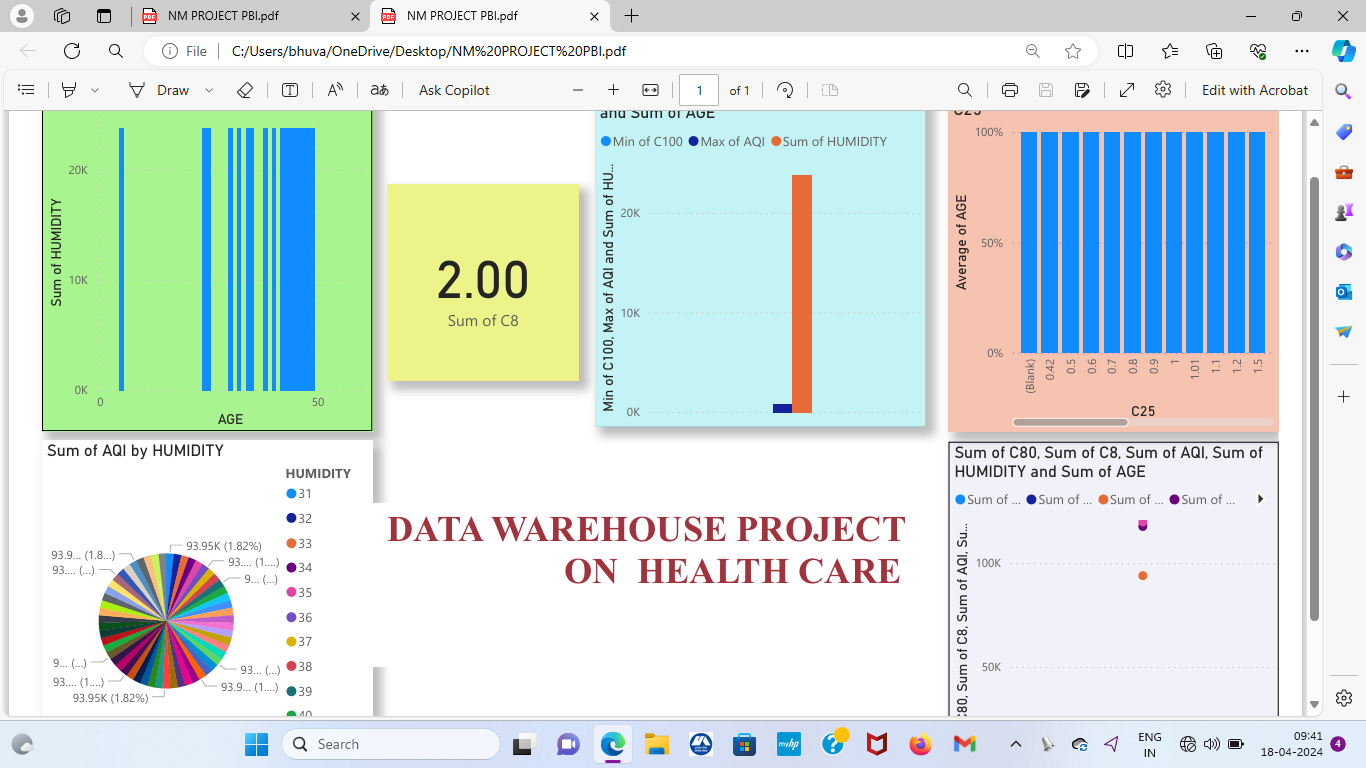
**Power BI Output:**

Choose a report view – choose pie chart in visualizations – choose a range of quality of air in the table – drag a column to the fields of x-axis and y-axis visualization—change the legends for colour change as your need.





**The Entire Power BI Representation:**



**Conclusion:**

* + In conclusion,undertaking for the healthcare sql project provided valuable insights with in the system to a real world datasets.
  + We can explore integrating sql with other techinques to further enhance health care data analysis capabilities.