Problem Description

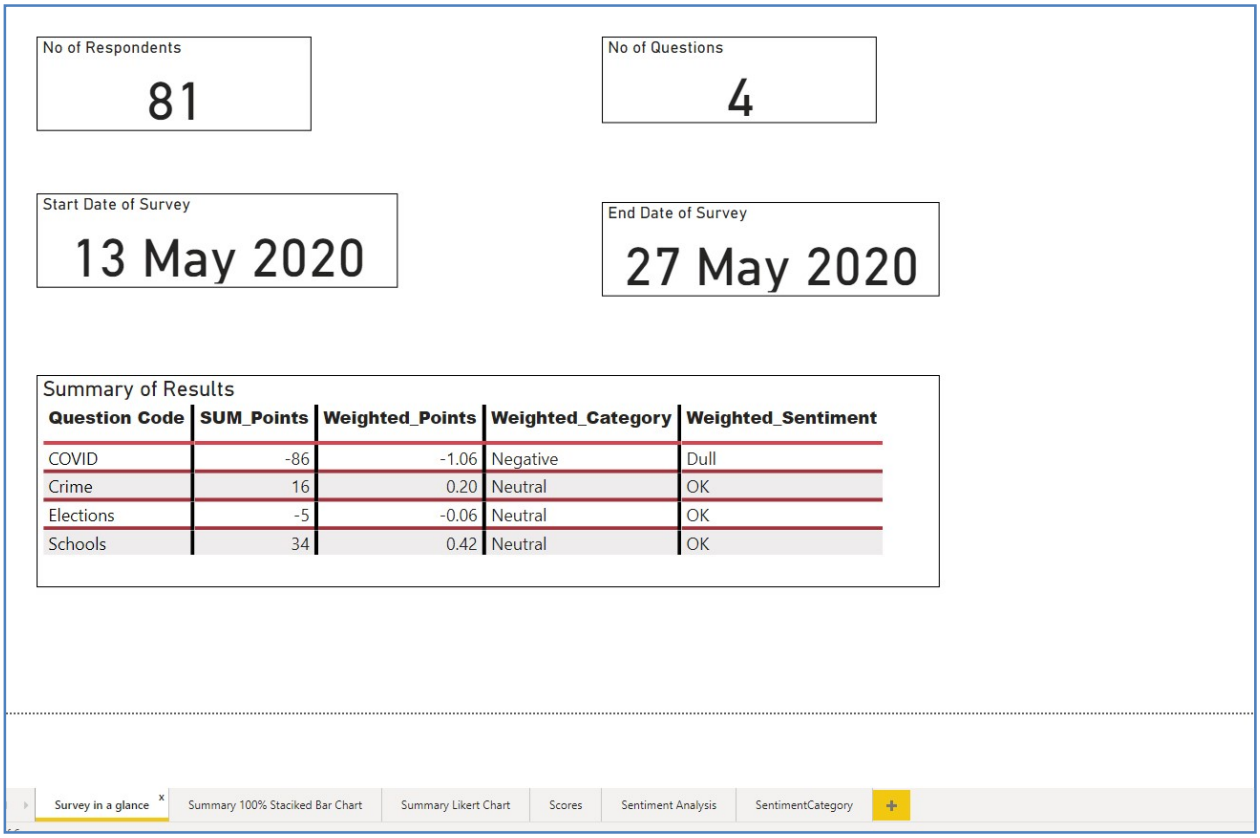
Case Study – Survey

Create a report to summarize the survey data. Survey data is based on the questions asked on four topics crime, COVID, school and elections. Create a report to compare the number of survey submissions and the score for each topic. Create a report to show the sentiments for each topic. Transform the data in proper format and create the required measures to generate the report.

Solution :

The following reports have been created.

1) Report 1 : - Survey in a glance



Question Code	SUM_Points	Weighted_Points	Weighted_Category	Weighted_Sentiment
COVID	-86	-1.06	Negative	Dull
Crime	16	0.20	Neutral	OK
Elections	-5	-0.06	Neutral	OK
Schools	34	0.42	Neutral	OK

Method

No of Respondents is a Measure which counts no of rows in Survey Table.
No of Questions is a Measure which counts no of rows of Question Table.
A new query Time&Duration is created from Start time and Completion time of Survey table.
The Start date and Completion date are extracted in new column Start Date and Completed Date .
Start Date of Survey is a Measure which is minimum value of Start Date column.
End Date of Survey is a Measure which is maximum value of Completion Date column/
The above measures are displayed using Card visual.

The Summary of Results is a Table visual of created Table Weighted _Score
The Question Code column is created by Enter Data and filling the Question Code manually.
The SUM_Points column is created by finding the SUM of the Points of Survey_Unpivot table by filtering the rows where the Question Code of Survey_Unpivot table and Question Code of Weighted _Score table are equal.
The Weighted_Points column is calculated by dividing the SUM_Points by No of respondents
The Weighted _Sentiment column is calculated using the IF statement considering the relationship between Points and Sentiment in SourceSentiment table.
The Weighted_Category column is derived from Weighted_Sentiment column using the SWITCH statement considering the relationship between Sentiment and Sentiment Category in SourceSentiment table.

Data Findings

The **Average Sentiment** considering all respondents is **OK** for Crime , Elections and Schools and **Dull** for COVID.
The Average Sentiment category considering all respondents is **Neutral** for Crime , Elections and Schools and **Negative** for COVID

2) Report 2 :- Summary 100 % Stacked Bar chart

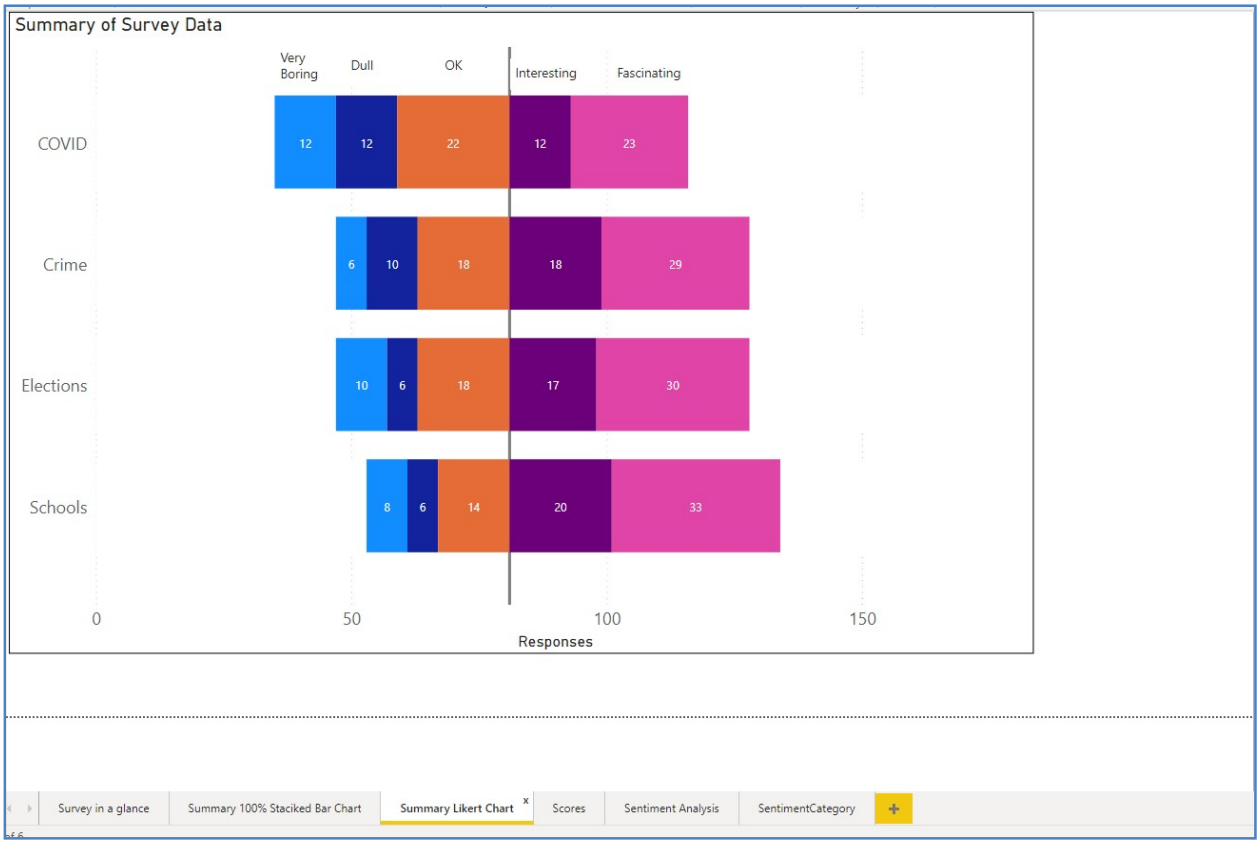
Method
The study of Survey table shows that for every id there are 6 columns of which 4 columns show the scores for each topic. To plot Summary we use 100 % Stacked bar chart visual for which we need data in two columns.

So the 4 column scores is converted to 2 column by using unpivot method. Hence new query Survey_Unpivot is created from Survey table. Also a conditional column of Points is created in the Survey_Unpivot table based on Points in SourceSentiment table.

Data Findings

Maximum % of No of responses are in **Fascinating** Sentiment for **COVID, Crime , Elections and Schools**
Minimum % of No of response are in **Very boring** Sentiment for **COVID and Crime** and in **Dull** Sentiment for **Elections and Schools**

3) Report 3 :- Summary Likert Chart



Method

In 100% stacked bar chart we can compare percentages of each scores for each topic but it does not give an idea about distribution of Positive , Neutral and Negative nature of Sentiment hence the Likert scale is used to arrange these percentages of Positive , Neutral + Negative on both sides of a central axis .

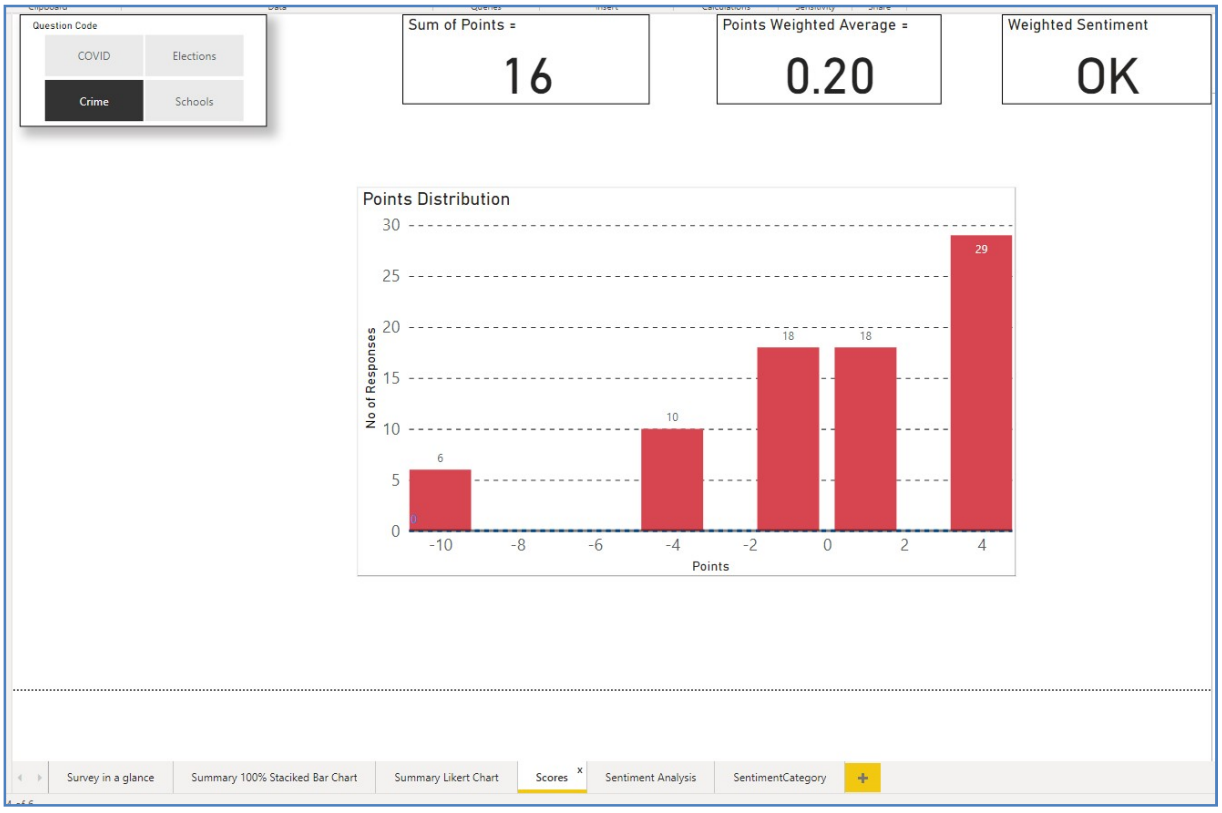
To arrange the Positive Sentiment bars of each topic along the same vertical axis buffers have to be added at both ends of the bars. The buffer added in the front is sum of the count of scores of Positive Sentiment ie Interesting and Fascinating and the buffer added at the end is the sum of the count of scores of Neutral + Negative Sentiment ie Very boring , Dull and OK .

Hence a new query Survey_Likert is created in which two scores 0 and 6 are added to the exiting scores of 1 to 5 with values as described above and a stacked bar chart visual is used to get Likert scale. Since the buffers ie scores 0 and 6 are only used to shift the bars they are made invisible by removing the fill colour in Data Colours Formatting.

The axis separating the Positive and Neutral + Negative Sentiment is created by using Analytics Constant Line feature.

Data Findings
In case of **Crime , Elections and Schools** the **Positive sentiment is more than Neutral + Negative sentiment.** In case of **COVID** the **Positive sentiment is less than Neutral + Negative sentiment.**

4) Report 4 :- Scores



Method

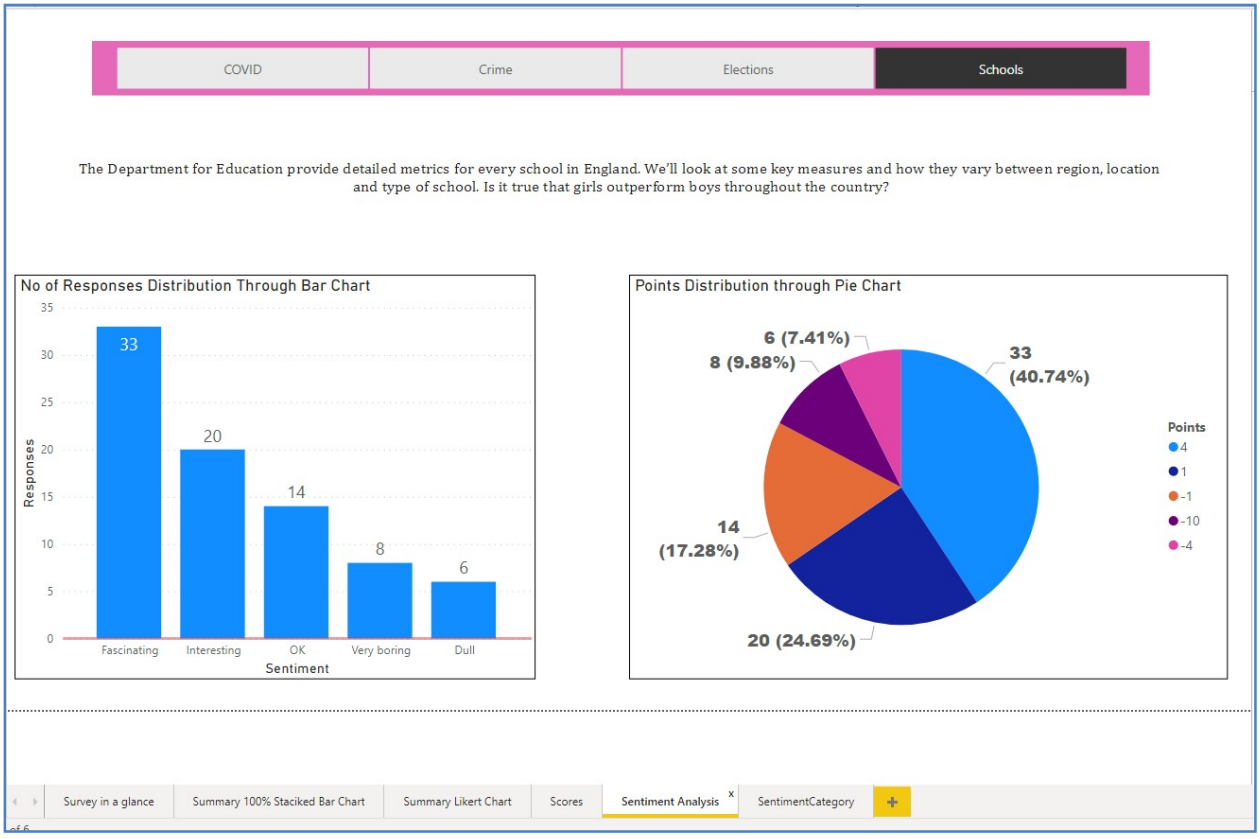
A slicer is used to select the question code.
The Sum of points , Points Weighted Average and Weighted Sentiment are displayed from Weighted_Score table using Card visual.

A Bar chart visual is used to plot the count of points for each Point type .

Data Findings
Even though the count of points for 4 point (Fascinating) and 1 point (Interesting) is very much higher than count for point -10 (Very boring) the overall Sum of points is lower due to high weight for Very boring sentiment as compared to other sentiment.

The overall Score for Crime topic is 0.2.
The overall sentiment for Crime topic is OK.

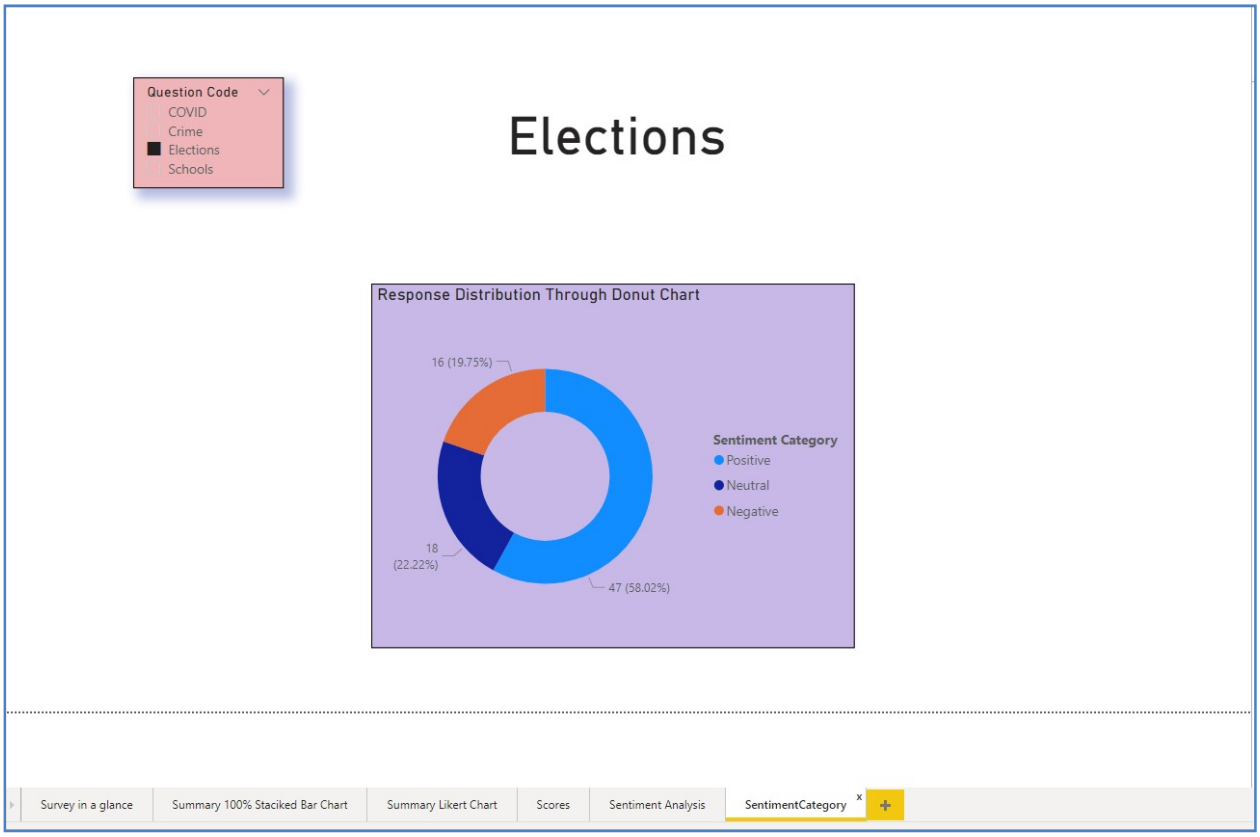
5) Report 5 :- Sentiment Analysis



Method
A slicer is used to select question code.
Based on selection the question description is displayed through Card visual.
The no of responses for each Sentiment category is displayed through bar chart visual.
The no of Points for each Points category is displayed through Pie chart visual.

Data Findings
For Schools topic the No of responses 33 is **highest** for **Fascinating** sentiment and the No of responses 6 is **lowest** for **Dull** sentiment. This can be seen in pie chart where **Point 4** has **highest percentage** 40.74 % and **Points -4** has **lowest percentage** (7.41 %)

6) Report 6 :- SentimentCategory



Method
A slicer is used to select question code.
The No of responses for each Sentiment Category is displayed through donut chart visual.

Data Findings
For Elections the No of responses 47 is **highest** for **Positive sentiment** and the No of responses 16 is **lowest** for **Negative sentiment**