Automated EDA on Visa dataset

Linkedin: https://www.linkedin.com/in/satya-nerurkar-9b0655190/
Github: https://github.com/SatyaNerurkar

In [1]:
```
# Install supporting libaries
!pip install pandas-profiling sweetviz --upgrade --quiet
```

In [2]:
```
# Import supporting libaries
import pandas as pd
from pandas_profiling import ProfileReport
import sweetviz as sv
import dtale
```

EDA using pandas profiling.

In [3]:
```
# EDA using pandas-profiling
profile = ProfileReport(pd.read_csv(".\Visadataset.csv"),
                        explorative=True)
```

In [4]:
```
# Saving results to a HTML file
profile.to_file("pandas_profiling_visa_report.html")

# Saving results to a JSON file
profile.to_file("pandas_profiling_visa_report.json")
```

In [5]:
```
# Displaying the report as a set of widgets
profile.to_widgets()
```

In [6]:
```
# Displaying report in notebook cell.
profile.to_notebook_iframe()
```

Overview

Dataset statistics

Number of variables	12
Number of observations	25480
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	14.1 MiB
Average record size in memory	581.9 B

Variable types

Categorical	6
Boolean	3
Numeric	3

Alerts

`case_id` has a high cardinality: 25480 distinct values	**High cardinality**
`case_id` is uniformly distributed	**Uniform**
`case_id` has unique values	**Unique**

EDA using Sweetviz

In [7]:
```
#EDA using Sweetviz
sweet_report = sv.analyze(pd.read_csv(".\Visadataset.csv"))
```

In [8]:
```
#Saving results to HTML file
sweet_report.show_html('sweet_visa_report.html')
```

```
Report sweet_visa_report.html was generated! NOTEBOOK/COLAB USERS: the web browser MAY not
pop up, regardless, the report IS saved in your notebook/colab files.
```

EDA using D-tale

In [9]:

```
# Assigning dataset to variable
df = pd.read_csv('.\Visadataset.csv')
dtale.show(df)
```

Out[9]: