# Word happiness report

```
In [1]: import numpy as np
        import pandas as pd
        import matplotlib.pyplot as plt
        %matplotlib inline
        plt.rcParams['figure.figsize'] = 10, 4
        plt.rcParams['font.family'] = 'sans-serif'
        pd.set_option('display.width', 500)
        pd.set_option('display.max_columns', 20)
        pd.set_option('display.notebook_repr_html', True)
        import seaborn as sns
        sns.set_style("whitegrid")
        sns.set_context("poster")
```

```
In [2]: df = pd.read_csv('2016.csv')
        df.head()
```

| | Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Denmark | Western Europe | 1 | 7.526 | 0.03411 | 1.44178 | 1.16374 | 0.79504 | 0.57004 | 0.39203 | 0.42571 | 2.12677 |
| 1 | Switzerland | Western Europe | 2 | 7.509 | 0.03806 | 1.52733 | 1.14524 | 0.86179 | 0.55477 | 0.41978 | 0.14106 | 2.09626 |
| 2 | Iceland | Western Europe | 3 | 7.501 | 0.04884 | 1.30232 | 1.18772 | 0.86061 | 0.54672 | 0.23179 | 0.48877 | 2.16320 |
| 3 | Norway | Western Europe | 4 | 7.498 | 0.03328 | 1.52537 | 1.18326 | 0.86091 | 0.56730 | 0.35776 | 0.36503 | 2.14474 |
| 4 | Finland | Western Europe | 5 | 7.413 | 0.03553 | 1.40598 | 1.13464 | 0.81091 | 0.55211 | 0.41372 | 0.23351 | 2.49204 |

```
In [3]: df.shape
```

```
Out[3]: (158, 12)
```

```
In [4]: df[df.Country.str.contains("Finland")]
```

| | Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | Finland | Western Europe | 5 | 7.413 | 0.03553 | 1.40598 | 1.13464 | 0.81091 | 0.55211 | 0.41372 | 0.23351 | 2.49204 |

```
In [5]: df.columns.tolist()
```

```
Out[5]: ['Country',
         'Region',
         'Happiness Rank',
         'Happiness Score',
         'Standard Error',
         'Economy (GDP per Capita)',
         'Family',
         'Health (Life Expectancy)',
         'Freedom',
         'Trust (Government Corruption)',
         'Generosity',
         'Dystopia Residual']
```

```
In [6]: df.dtypes.sort_values
```

```
Out[6]: Country                        object
        Region                         object
        Happiness Rank                  int64
        Happiness Score                float64
        Standard Error                 float64
        Economy (GDP per Capita)       float64
        Family                         float64
        Health (Life Expectancy)       float64
        Freedom                        float64
        Trust (Government Corruption)  float64
        Generosity                     float64
        Dystopia Residual              float64
        dtype: object
```

```
In [7]: df[df.Region.isnull]
```

| | Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 21 | Luxembourg | NaN | 22 | 6.946 | 0.04053 | 1.609 | 1.1305 | 0.8123 | 0.5351 | 0.3708 | 0.2467 | 2.1626 |
| 22 | Mexico | NaN | 23 | 6.595 | 0.05499 | 1.30342 | 1.07055 | 0.74393 | 0.30177 | 0.07223 | 0.36051 | 2.19719 |
| 23 | Chile | NaN | 24 | 6.505 | 0.05348 | 1.20567 | 1.10062 | 0.79179 | 0.44578 | 0.08466 | 0.12223 | 2.32358 |
| 24 | Brazil | NaN | 25 | 6.498 | 0.05465 | 1.05351 | 1.00537 | 0.69837 | 0.43643 | 0.09049 | 0.16723 | 2.49160 |
| 25 | Argentina | NaN | 26 | 6.488 | 0.05884 | 1.09568 | 1.05351 | 0.74796 | 0.35775 | 0.06148 | 0.09011 | 2.63933 |
| 26 | Uruguay | NaN | 27 | 6.474 | 0.05814 | 1.26017 | 1.05351 | 0.76576 | 0.43482 | 0.06982 | 0.07004 | 2.37152 |
| 27 | Japan | NaN | 28 | 6.465 | 0.04732 | 1.39488 | 0.94284 | 0.94284 | 0.26666 | 0.32333 | 0.12744 | 2.04626 |
| 28 | Bahrain | NaN | 29 | 6.449 | 0.06174 | 1.42727 | 0.92039 | 0.67744 | 0.42908 | 0.12575 | 0.18772 | 2.24089 |
| 29 | Slovakia | NaN | 30 | 6.411 | 0.05348 | 1.26404 | 1.08772 | 0.73086 | 0.35486 | 0.02832 | 0.09403 | 2.41863 |

```
In [8]: df[df.Region.isnull].Region.fillna('Latin America and Caribbean')
```

```
Out[8]: 21    Latin America and Caribbean
        22    Latin America and Caribbean
        23    Latin America and Caribbean
        24    Latin America and Caribbean
        25    Latin America and Caribbean
        26    Latin America and Caribbean
        27    Latin America and Caribbean
        28    Latin America and Caribbean
        29    Latin America and Caribbean
        Name: Region, dtype: object
```

```
In [9]: df[df.Region.isnull].Region = df[df.Region.isnull].Region.fillna('Latin America and Caribbean')
```

```
In [10]: df[df.Region.isnull].head(10)
```

| | Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 21 | Luxembourg | NaN | 22 | 6.946 | 0.04053 | 1.609 | 1.1305 | 0.8123 | 0.5351 | 0.3708 | 0.2467 | 2.1626 |
| 22 | Mexico | NaN | 23 | 6.595 | 0.05499 | 1.30342 | 1.07055 | 0.74393 | 0.30177 | 0.07223 | 0.36051 | 2.19719 |
| 23 | Chile | NaN | 24 | 6.505 | 0.05348 | 1.20567 | 1.10062 | 0.79179 | 0.44578 | 0.08466 | 0.12223 | 2.32358 |
| 24 | Brazil | NaN | 25 | 6.498 | 0.05465 | 1.05351 | 1.00537 | 0.69837 | 0.43643 | 0.09049 | 0.16723 | 2.49160 |
| 25 | Argentina | NaN | 26 | 6.488 | 0.05884 | 1.09568 | 1.05351 | 0.74796 | 0.35775 | 0.06148 | 0.09011 | 2.63933 |
| 26 | Uruguay | NaN | 27 | 6.474 | 0.05814 | 1.26017 | 1.05351 | 0.76576 | 0.43482 | 0.06982 | 0.07004 | 2.37152 |
| 27 | Japan | NaN | 28 | 6.465 | 0.04732 | 1.39488 | 0.94284 | 0.94284 | 0.26666 | 0.32333 | 0.12744 | 2.04626 |
| 28 | Bahrain | NaN | 29 | 6.449 | 0.06174 | 1.42727 | 0.92039 | 0.67744 | 0.42908 | 0.12575 | 0.18772 | 2.24089 |
| 29 | Slovakia | NaN | 30 | 6.411 | 0.05348 | 1.26404 | 1.08772 | 0.73086 | 0.35486 | 0.02832 | 0.09403 | 2.41863 |

```
In [11]: df[df.Region.isnull].Region.fillna('Latin America and Caribbean')
```

```
Out[11]: 21    Latin America and Caribbean
         22    Latin America and Caribbean
         23    Latin America and Caribbean
         24    Latin America and Caribbean
         25    Latin America and Caribbean
         26    Latin America and Caribbean
         27    Latin America and Caribbean
         28    Latin America and Caribbean
         29    Latin America and Caribbean
         Name: Region, dtype: object
```

```
In [12]: df[df.Region.isnull].Region = df[df.Region.isnull].Region.fillna('Latin America and Caribbean')
```

```
In [13]: df[df.Region.isnull].Region
```

```
Out[13]: Series([], Name: Region, dtype: object)
```

```
In [14]: df[df.Country.isnull]
```

| | Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 158 | NaN | Sub-Saharan Africa | 159 | 3.083 | 0.07315 | 0.28123 | 0.67042 | 0.15775 | 0.35386 | 0.06773 | 0.20246 | 1.83772 |

```
In [15]: df.shape
```

```
Out[15]: (159, 12)
```

```
In [16]: df.dropna(inplace=True)
         df.shape
```

```
Out[16]: (158, 12)
```

```
In [17]: df[df.Country.isnull]
```

```
Out[17]: Series([], Dtype: bool)
```

```
In [18]: df[df.Region.isnull].Region
```

```
Out[18]: Series([], Name: Region, dtype: object)
```

```
In [19]: df[df.Region.isnull].Region = df[df.Region.isnull].Region.fillna('Sub-Saharan Africa')
         df[df.Country.isnull].Region
```

```
Out[19]: 158    Sub-Saharan Africa
         Name: Region, dtype: object
```

```
In [20]: df[df.Region.isnull].Region
```

```
Out[20]: Series([], Name: Region, dtype: object)
```

```
In [21]: df[df.Country.isnull].Region = df[df.Country.isnull].Region.fillna('Sub-Saharan Africa')
         df[df.Country.isnull].Region
```

```
Out[21]: 158    Sub-Saharan Africa
         Name: Region, dtype: object
```

```
In [22]: df[df.Country.isnull].Region = df[df.Country.isnull].Region.fillna('Sub-Saharan Africa')
         df[df.Country.isnull].Region
```

```
Out[22]: 158    Sub-Saharan Africa
         Name: Region, dtype: object
```

```
In [23]: df[df.Country.isnull].Region = df[df.Country.isnull].Region.fillna('Sub-Saharan Africa')
         df[df.Country.isnull].Region
```

```
Out[23]: 158    Sub-Saharan Africa
         Name: Region, dtype: object
```

```
In
```