

World Happiness

Dataset

Marian Dumitru Danci

The Data

	Country	Happiness Rank	Happiness Score	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Generosity	Dystopia Residual	Year
0	Switzerland	1	7587	1.39651	1.34951	0.94143	0.66557	0.41978	0.29678	2.51738	2015
1	Iceland	2	7561	1.30232	1.40223	0.94784	0.62877	0.14145	0.43630	2.70201	2015
2	Denmark	3	7527	1.32548	1.36058	0.87464	0.64938	0.48357	0.34139	2.49204	2015
3	Norway	4	7522	1.45911	1.33095	0.88521	0.66973	0.36503	0.34699	2.46531	2015
4	Canada	5	7427	1.32629	1.32261	0.90563	0.63297	0.32957	0.45811	2.45176	2015

```
RangeIndex: 470 entries, 0 to 469
```

```
Data columns (total 11 columns):
```

```
Country                470 non-null object
Happiness Rank         470 non-null int64
Happiness Score        470 non-null int64
Economy (GDP per Capita) 467 non-null float64
Family                 467 non-null float64
Health (Life Expectancy) 467 non-null float64
Freedom                 465 non-null float64
Trust (Government Corruption) 465 non-null float64
Generosity              467 non-null float64
Dystopia Residual        470 non-null float64
Year                    470 non-null int64
```

```
dtypes: float64(7), int64(3), object(1)
```

The Data

	Happiness Rank	Happiness Score	Economy (GDP per Capita)	Family	Health (Life Expectancy)	Freedom	Trust (Government Corruption)	Generosity	Dystopia Residual	Year
count	470.000000	470.000000	467.000000	467.000000	467.000000	465.000000	465.000000	467.000000	470.000000	470.000000
mean	78.829787	5370.712766	0.933791	0.998213	0.583695	0.406673	0.135663	0.243799	2.091161	2015.995745
std	45.281408	1137.001666	0.410176	0.307183	0.236366	0.146342	0.110766	0.130516	0.564178	0.815615
min	1.000000	2693.000000	0.015300	0.104190	0.005565	0.005890	0.002270	0.001990	0.328580	2015.000000
25%	40.000000	4509.000000	0.622300	0.801848	0.406966	0.306580	0.060750	0.154854	1.737975	2015.000000
50%	79.000000	5282.500000	0.996020	1.029120	0.631040	0.422840	0.100107	0.224151	2.094640	2016.000000
75%	118.000000	6233.750000	1.256767	1.230947	0.771255	0.518631	0.172502	0.316211	2.453263	2017.000000
max	158.000000	7587.000000	1.870766	1.610574	1.025250	0.669730	0.551910	0.838075	3.837720	2017.000000

Objectives



Exploratory Data Analysis

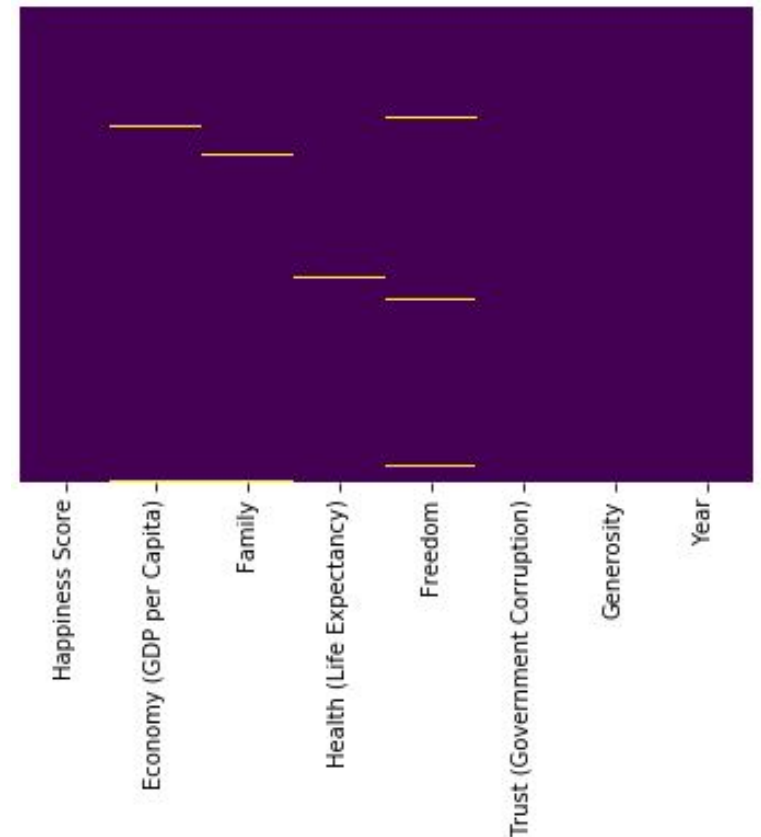
- Data Cleaning
- Missing Data

```
Int64Index: 452 entries, 0 to 468
```

```
Data columns (total 8 columns):
```

Happiness Score	452	non-null	int64
Economy (GDP per Capita)	452	non-null	float64
Family	452	non-null	float64
Health (Life Expectancy)	452	non-null	float64
Freedom	452	non-null	float64
Trust (Government Corruption)	452	non-null	float64
Generosity	452	non-null	float64
Year	452	non-null	int64

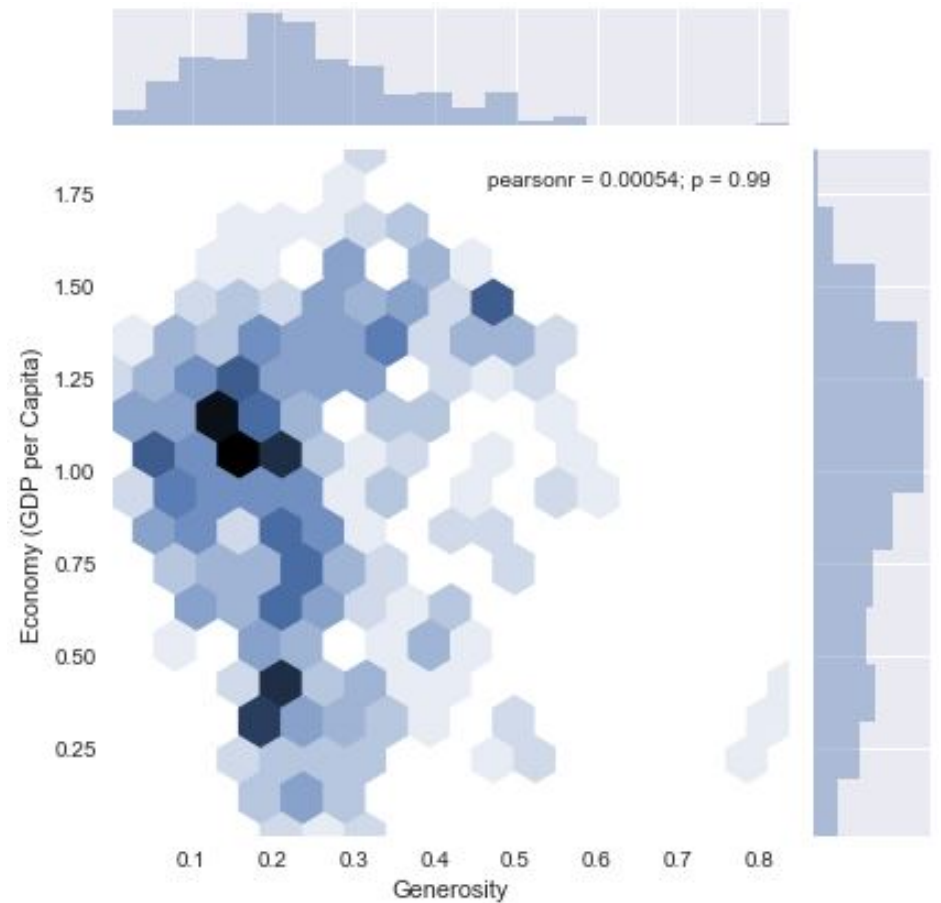
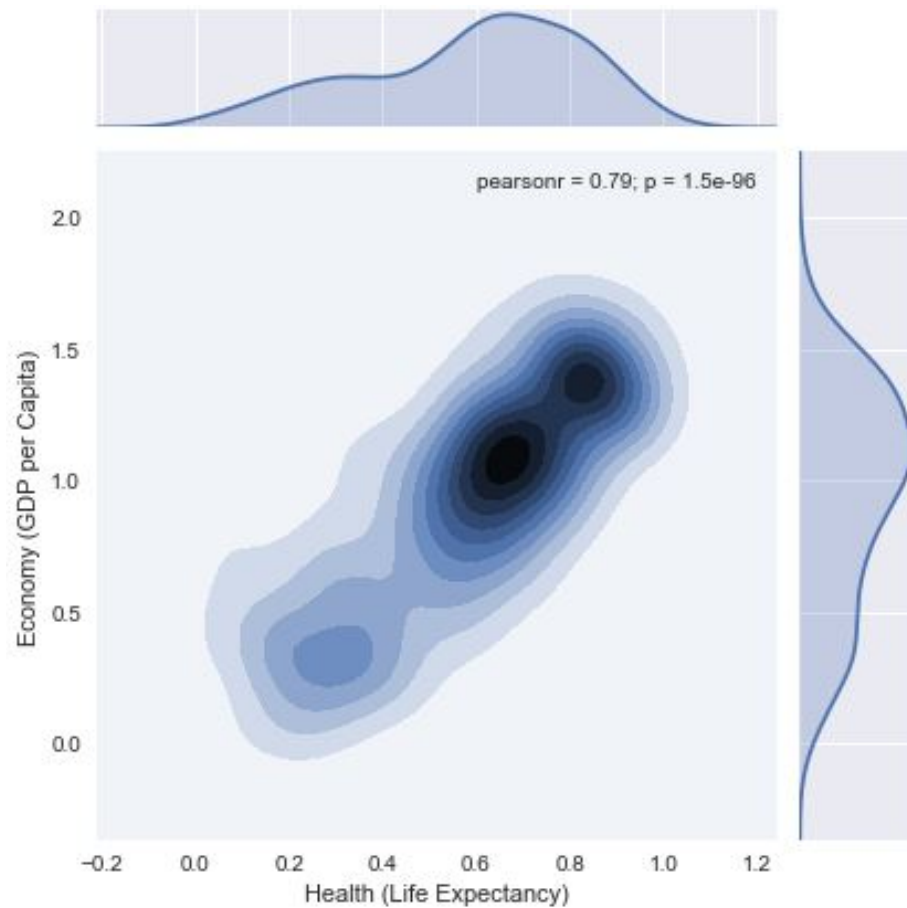
```
dtypes: float64(6), int64(2)
```



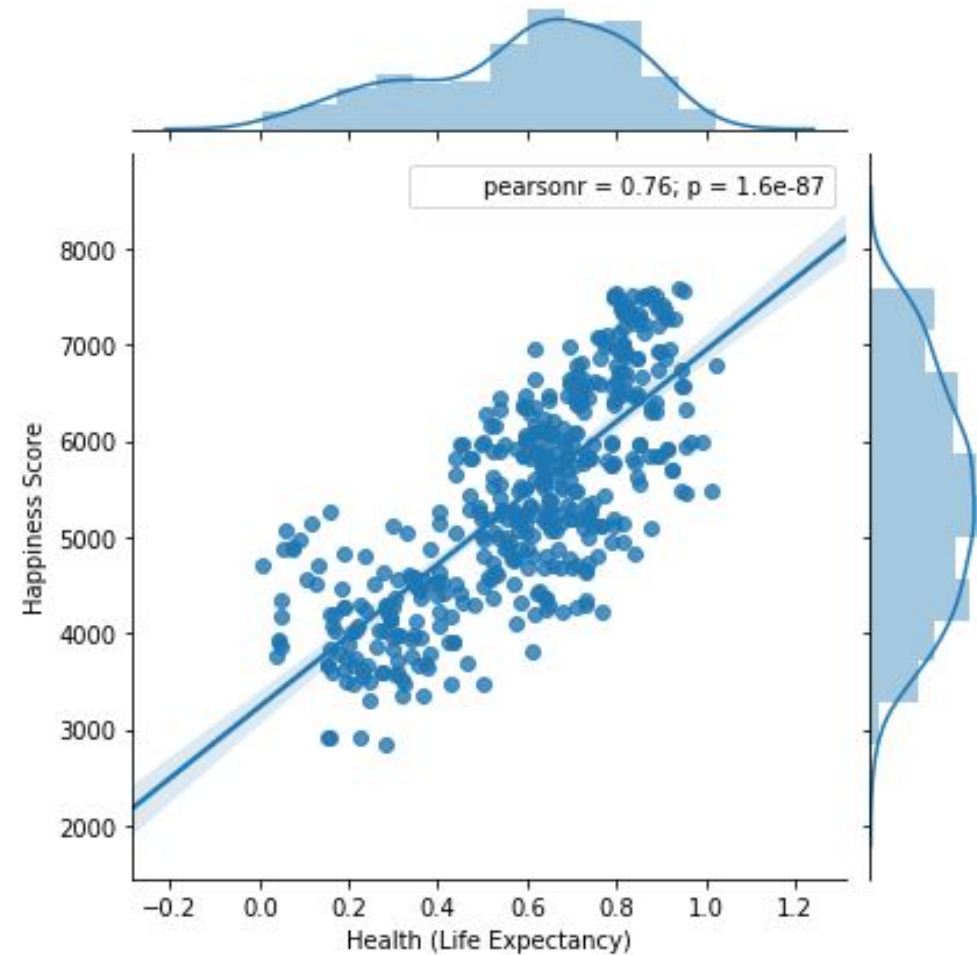
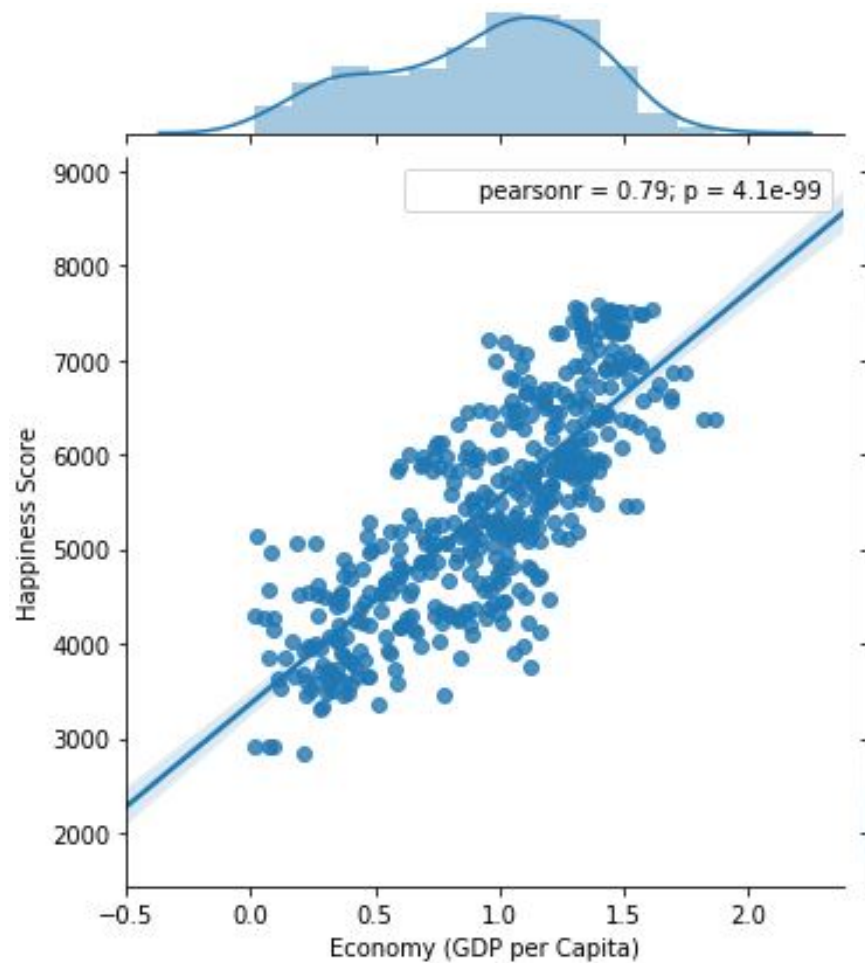
Exploratory Data Analysis



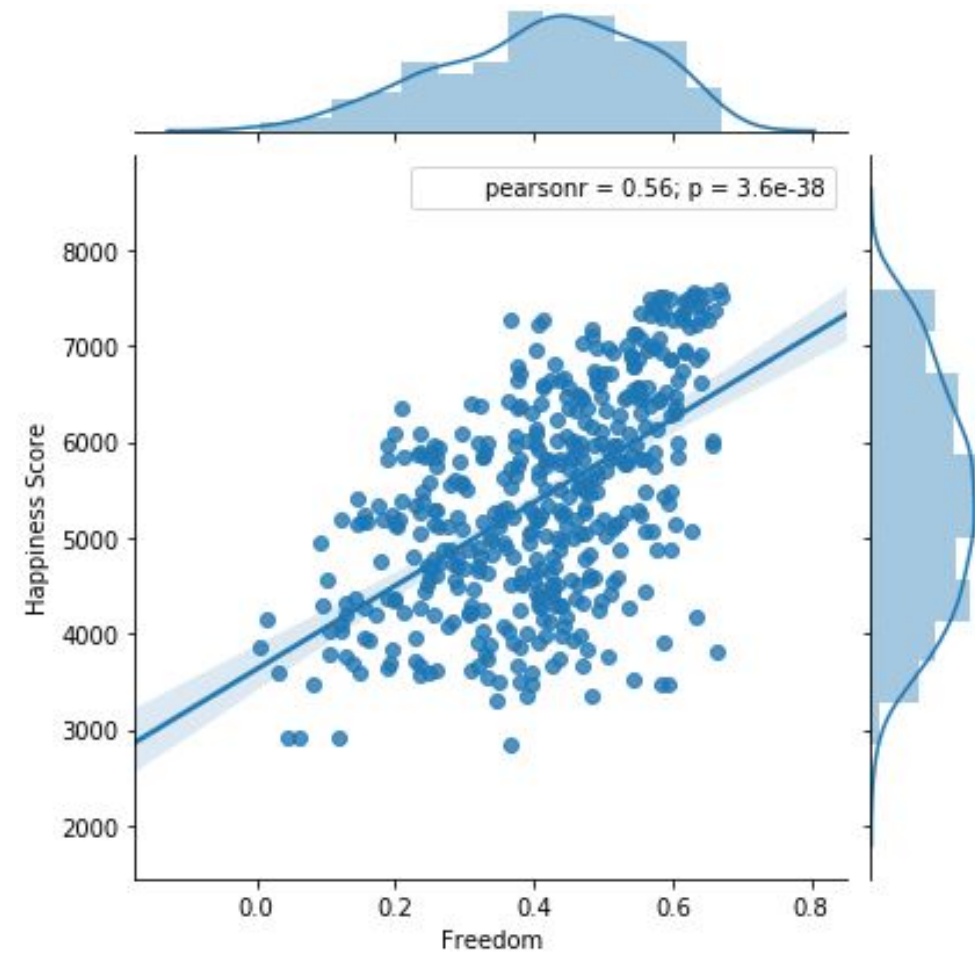
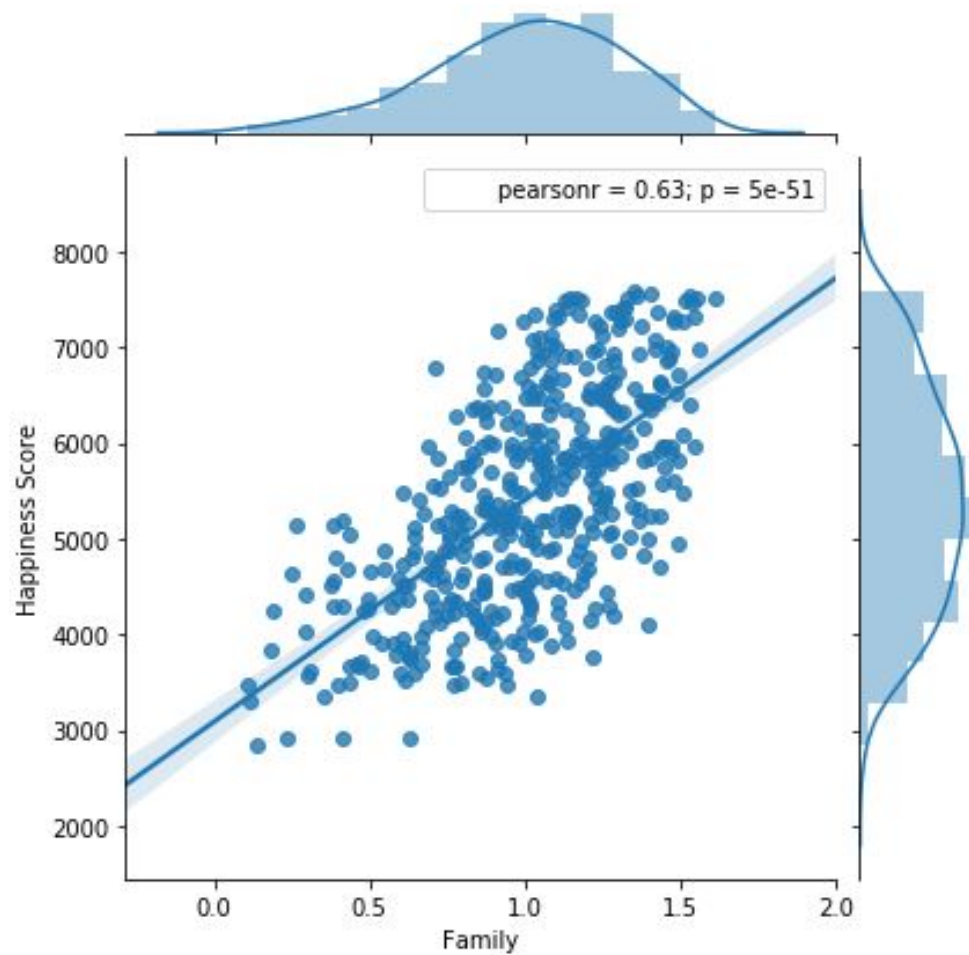
Exploratory Data Analysis



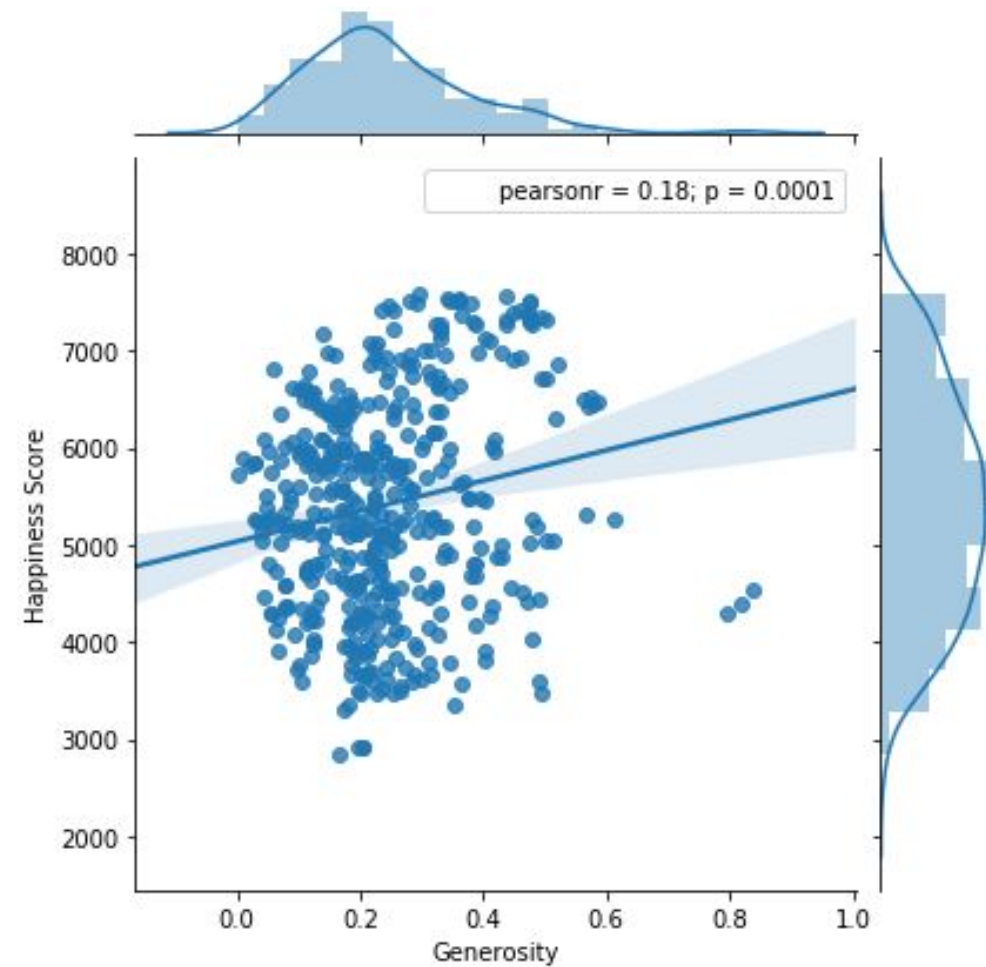
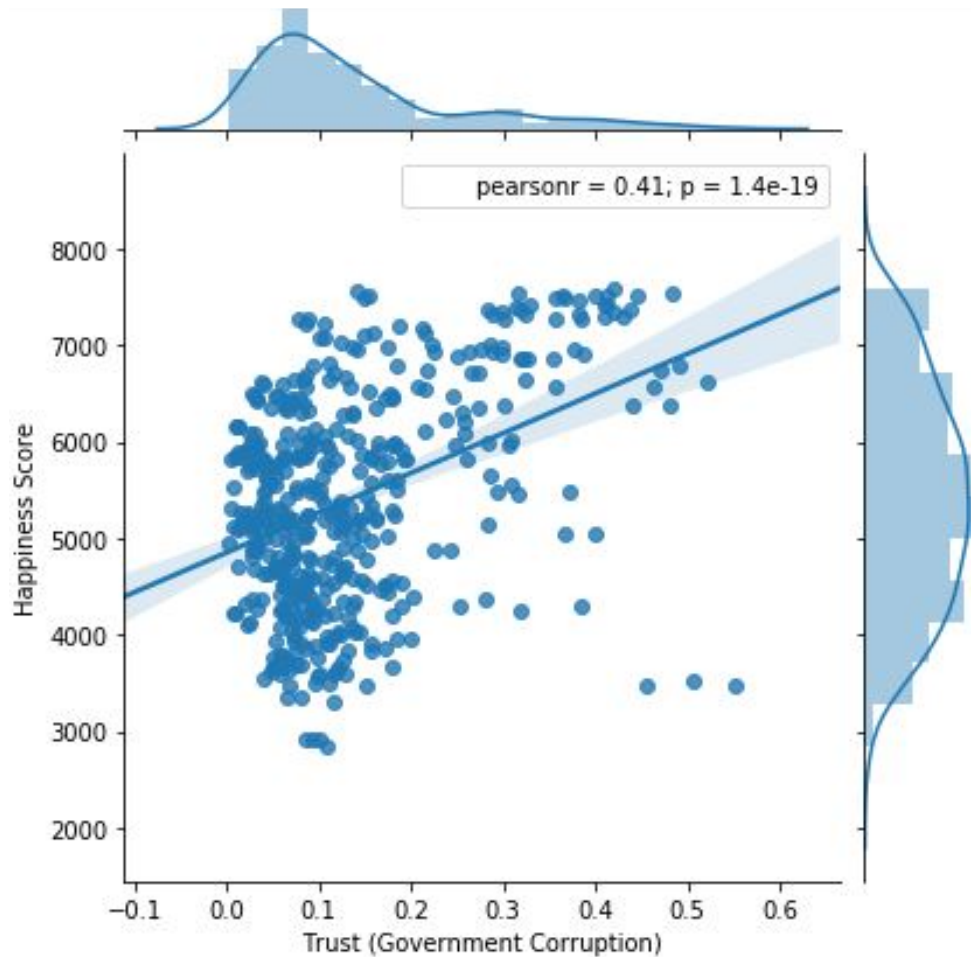
Exploratory Data Analysis



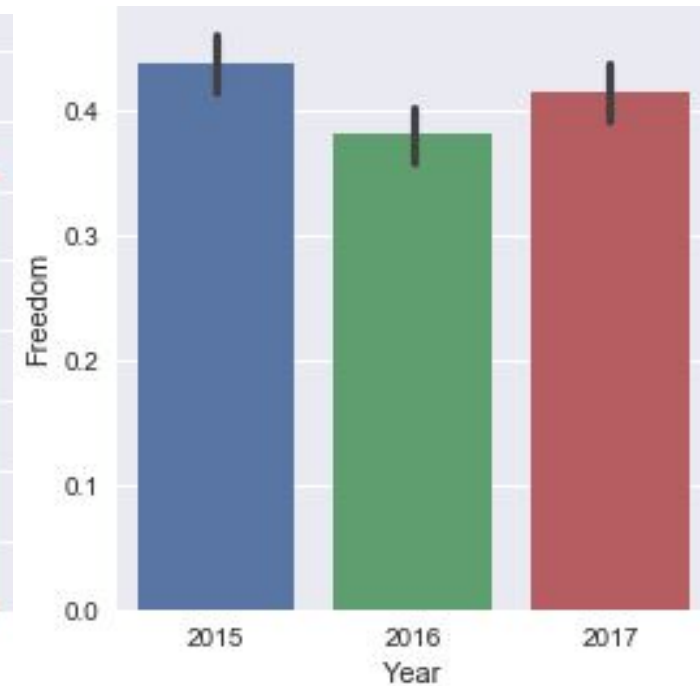
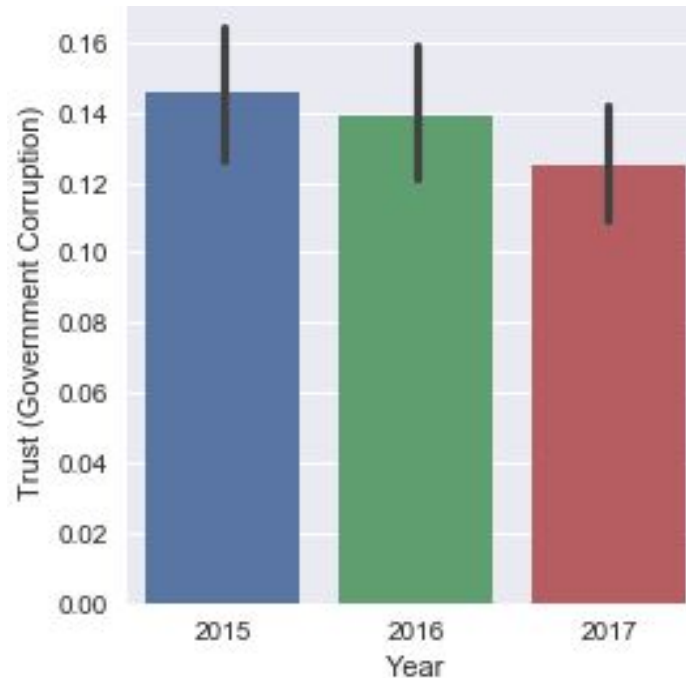
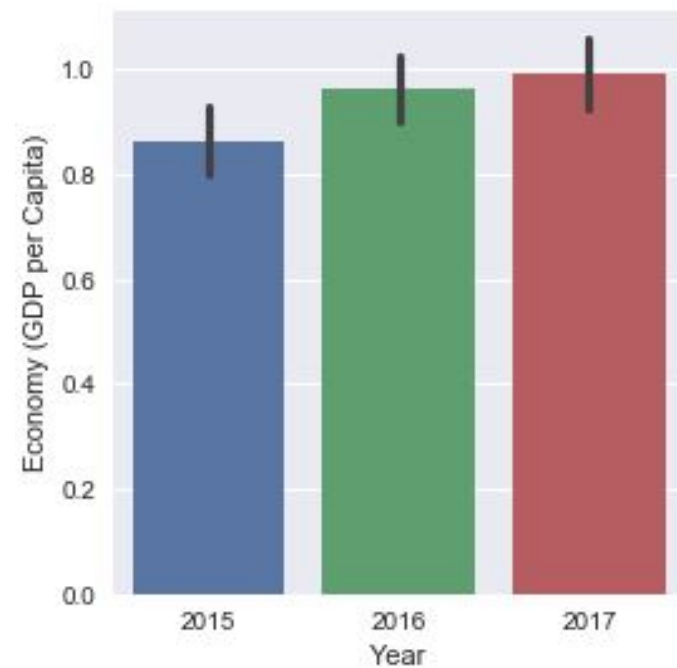
Exploratory Data Analysis



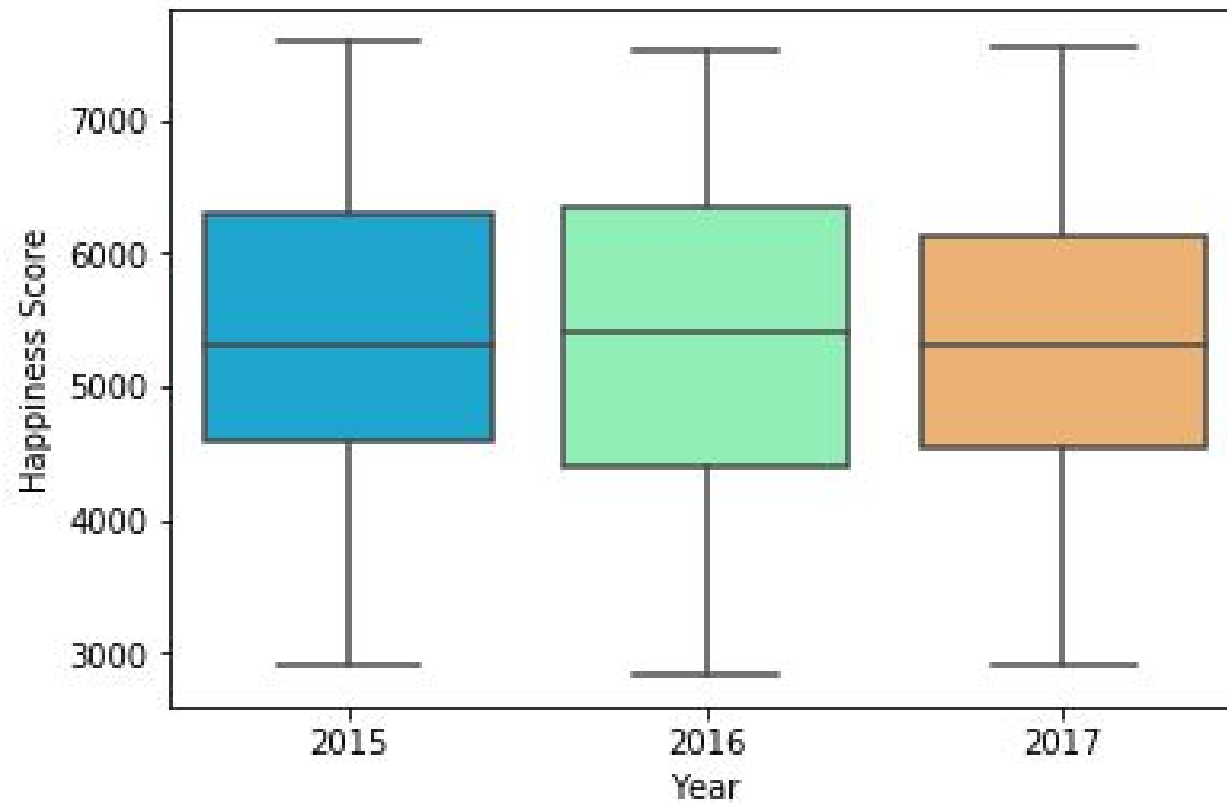
Exploratory Data Analysis



Exploratory Data Analysis



Exploratory Data Analysis

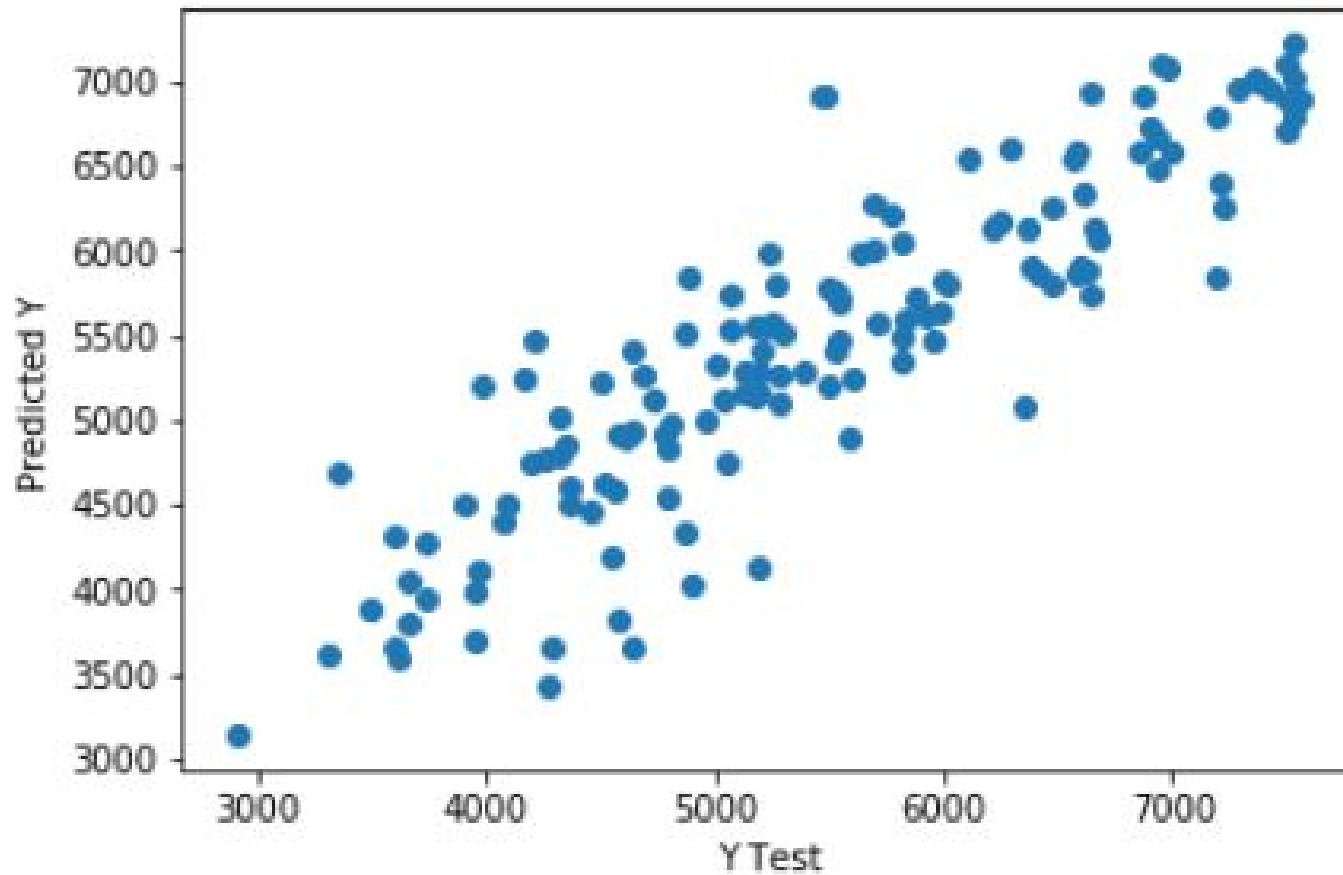


Linear Regression Model

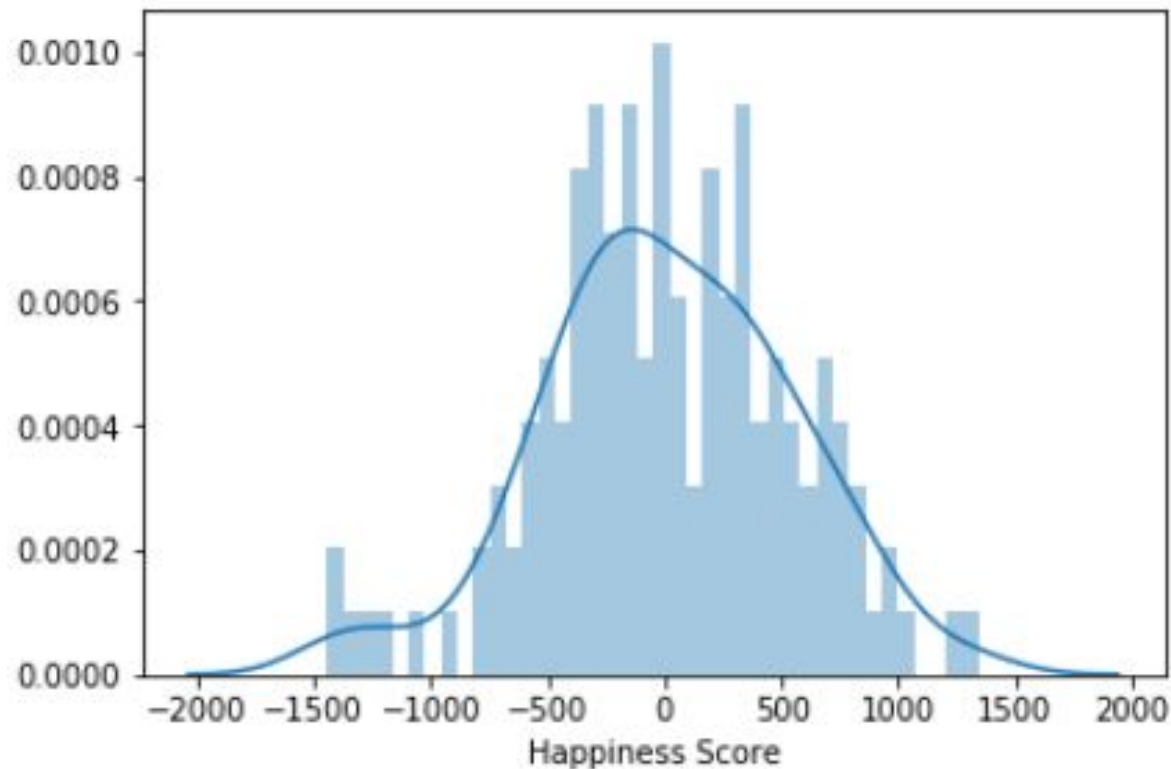
- Features to train: X
- Target variable: y

```
X = WH[['Economy (GDP per Capita)', 'Family', 'Health (Life Expectancy)',  
        'Freedom', 'Generosity', 'Trust (Government Corruption)']]  
  
y = WH['Happiness Score']
```

Linear Regression Model



Linear Regression Model



- Mean Absolute Error (MAE): 424.419
- Mean Squared Error (MSE): 286165.191
- Root Mean Squared Error (RMSE): 534.944

Conclusion

	Coefficient
Economy (GDP per Capita)	1049.761702
Family	619.666244
Health (Life Expectancy)	1203.698239
Freedom	1642.653014
Generosity	357.603414
Trust (Government Corruption)	577.756056



Questions?