employeeEDA

January 23, 2018

0.1 Employee EDA

Matt Oehler

In [1]:
```
#Import LIbraries
#for data manipulation and operations
import pandas as pd

#for plotting
from matplotlib import pyplot as plt
%matplotlib inline
import seaborn as sns
```

In [14]:
```
#read in the data
dat = pd.read_table('employee.txt',delim_whitespace=True)
print(dat.head())
print(dat.shape)
```

```
ID  Age  Tenure  WellBeing  JobSat  JobPerf      IQ
0   1   40      10       7.92    8.33     5.63  105.93
1   2   53      14       6.21    4.99      NaN   92.69
2   3   46      10        NaN    7.39     7.30  106.76
3   4   37       8       6.51     NaN     5.02   94.27
4   5   44       9        NaN    5.05      NaN  107.43
(480, 7)
```

In [10]:
```
dat.describe()
```

Out[10]:

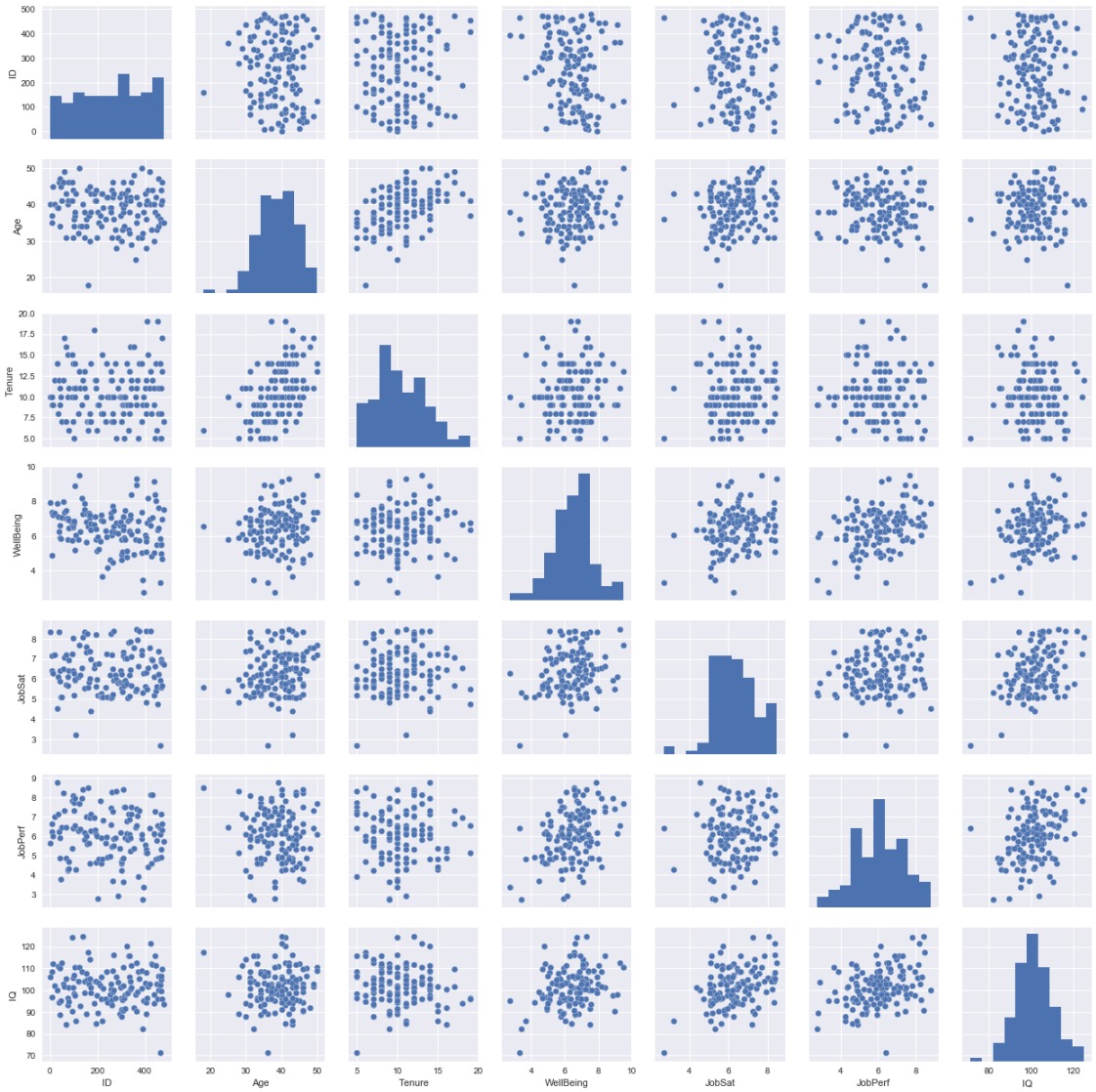
	ID	Age	Tenure	WellBeing	JobSat	JobPerf	IQ
count	480.000000	480.000000	480.000000	320.000000	320.000000	416.000000	480.000000
mean	240.500000	37.947917	10.054167	6.269687	5.953219	6.073846	100.114646
std	138.708327	5.382198	3.123282	1.198116	1.239017	1.266946	8.425031
min	1.000000	18.000000	1.000000	2.750000	2.670000	2.730000	71.070000
25%	120.750000	34.000000	8.000000	5.455000	5.060000	5.292500	94.515000
50%	240.500000	38.000000	10.000000	6.320000	5.940000	6.070000	99.915000
75%	360.250000	42.000000	12.000000	7.102500	6.837500	6.930000	
max	480.000000	53.000000	21.000000	9.500000	9.370000	9.580000	

```
          75%    105.515000
          max    124.810000
```

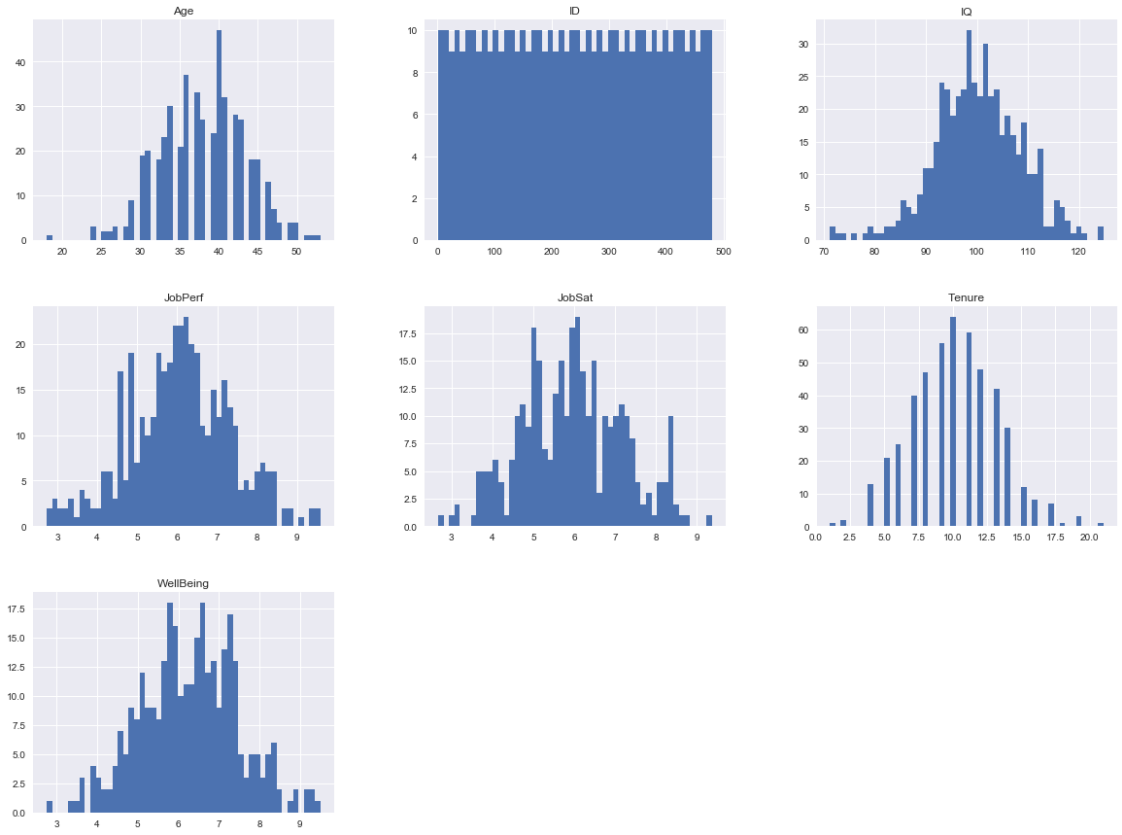
In [12]:
```
dat.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 480 entries, 0 to 479
Data columns (total 7 columns):
ID            480 non-null int64
Age           480 non-null int64
Tenure        480 non-null int64
WellBeing     320 non-null float64
JobSat        320 non-null float64
JobPerf       416 non-null float64
IQ            480 non-null float64
dtypes: float64(4), int64(3)
memory usage: 26.3 KB
```

In [8]:
```
sns.pairplot(dat.dropna(axis=0,how='any'))
sns.despine()
```



In [15]:
```
#make histograms of all the variables
dat.hist(bins=50,figsize=(20,15))
plt.show()
```



In [ ]:

In [ ]: