Regression task: The given dataset has 1197 rows and 15 columns.

pandas: Importing databases

```
[144] import numpy as np
      import pandas as pd
      import matplotlib.pyplot as plt
      import seaborn as sns
```

Reading the regression dataset i.e. garments_worker_productivity.csv

```
[145] df_pr = pd.read_csv("garments_worker_productivity.csv")
```

displaying dataset i.e. the csv file

```
df_pr
```

	date	quarter	department	day	team	targeted_productivity	smv	wip	over_time	incentive	idle_time	idle_men	no_of_style_change	no_of_workers
0	1/1/2015	Quarter1	sewing	Thursday	8	0.80	26.16	1108.0	7080	98	0.0	0	0	59.0
1	1/1/2015	Quarter1	finishing	Thursday	1	0.75	3.94	NaN	960	0	0.0	0	0	8.0
2	1/1/2015	Quarter1	sewing	Thursday	11	0.80	11.41	968.0	3660	50	0.0	0	0	30.5
3	1/1/2015	Quarter1	sewing	Thursday	12	0.80	11.41	968.0	3660	50	0.0	0	0	30.5
4	1/1/2015	Quarter1	sewing	Thursday	6	0.80	25.90	1170.0	1920	50	0.0	0	0	56.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1192	3/11/2015	Quarter2	finishing	Wednesday	10	0.75	2.90	NaN	960	0	0.0	0	0	8.0
1193	3/11/2015	Quarter2	finishing	Wednesday	8	0.70	3.90	NaN	960	0	0.0	0	0	8.0
1194	3/11/2015	Quarter2	finishing	Wednesday	7	0.65	3.90	NaN	960	0	0.0	0	0	8.0
1195	3/11/2015	Quarter2	finishing	Wednesday	9	0.75	2.90	NaN	1800	0	0.0	0	0	15.0
1196	3/11/2015	Quarter2	finishing	Wednesday	6	0.70	2.90	NaN	720	0	0.0	0	0	6.0

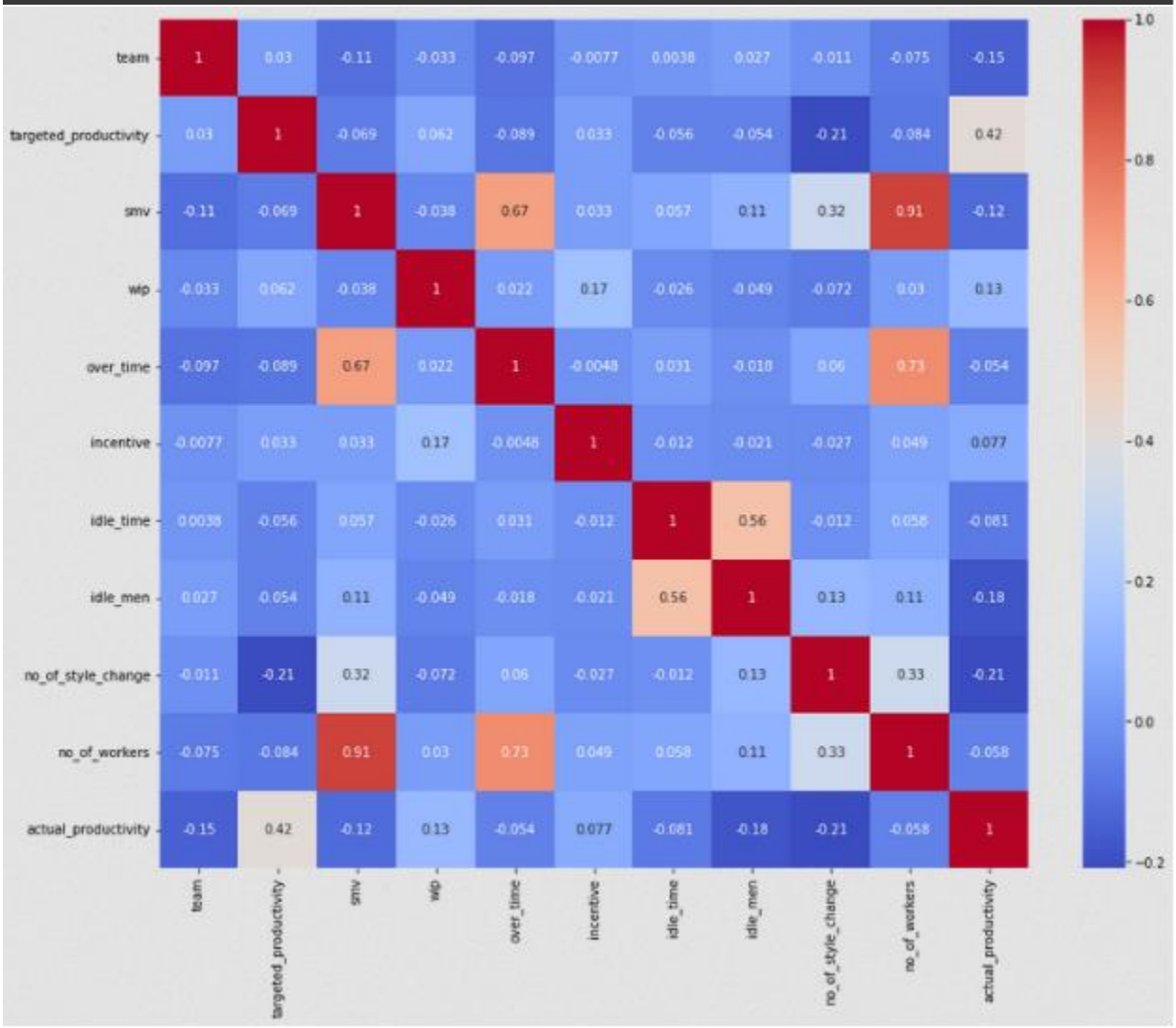
1197 rows × 15 columns

```
df_pr.shape
```

```
(1197, 15)
```

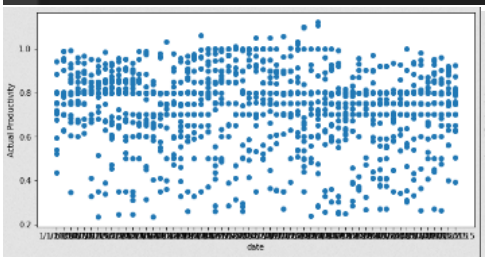
Heatmap

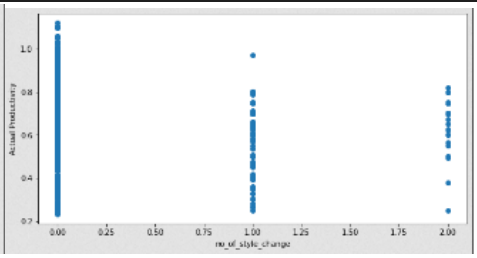
```
[149] plt.figure(figsize = (15,12))
      sns.heatmap(data=df_pr.corr(),annot=True, cmap="coolwarm")
```

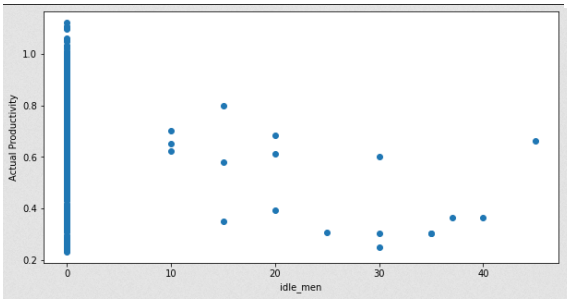


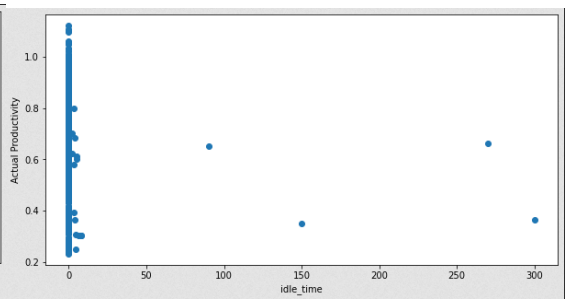
Scatter plots

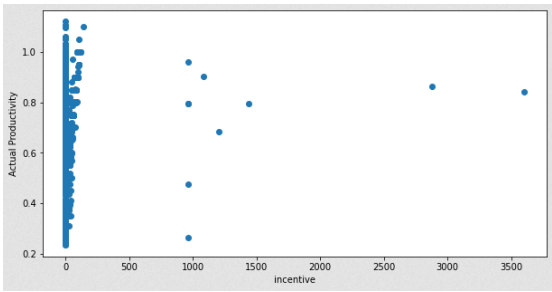
```
[150] for i in range(df_pr.shape[1]):
        plt.figure(figsize = (10,5))
        #if(i == 0):
        plt.scatter(df_pr.iloc[:, i], df_pr['actual_productivity'])
        #else:
          #plt.scatter(df_pr.iloc[:, i], df_pr['actual_productivity'])
        plt.xlabel(df_pr.columns[i])
        plt.ylabel("Actual Productivity")
```

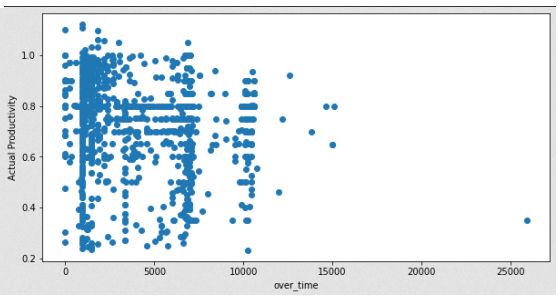


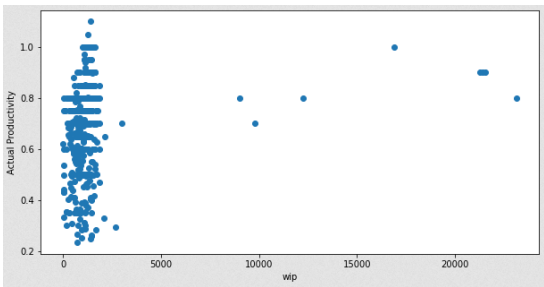


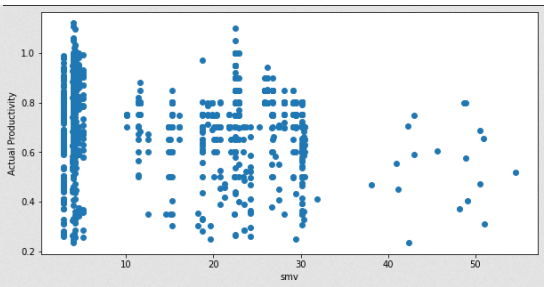


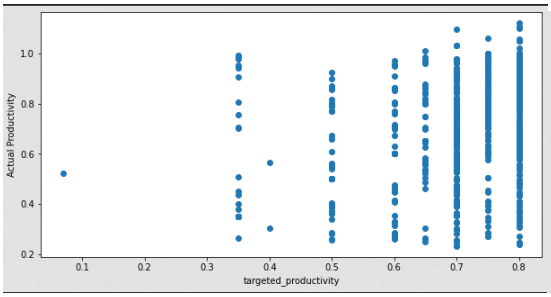


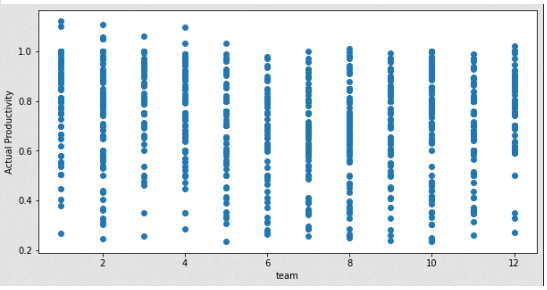


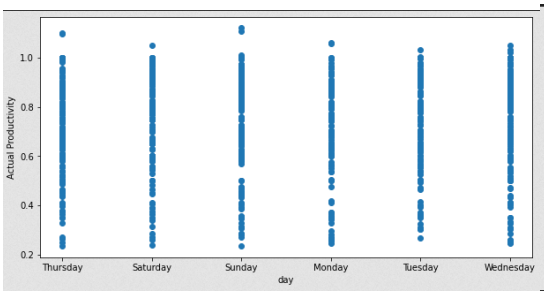


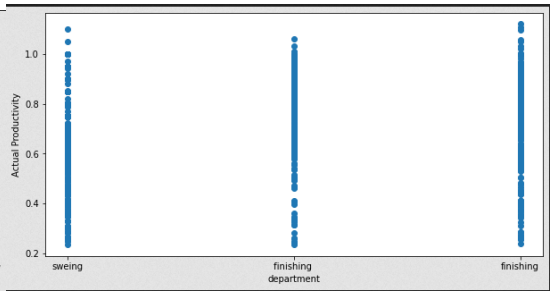


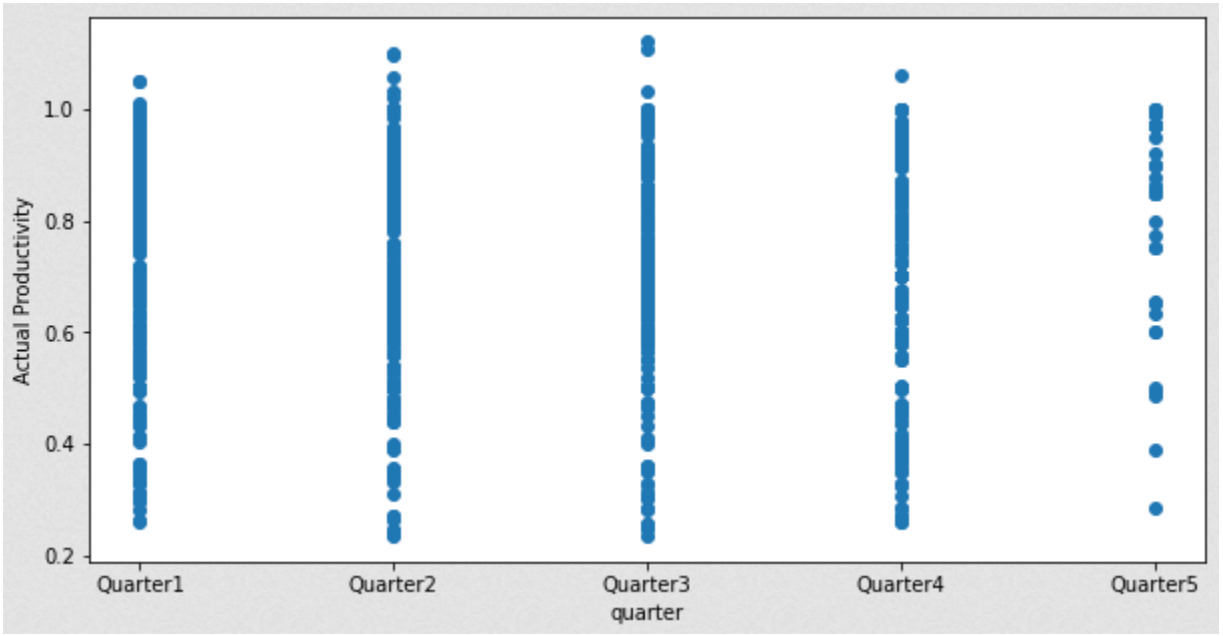



Actual Productivity
targeted_productivity


Actual Productivity
team


Actual Productivity
Thursday
Saturday
Sunday
Monday
Tuesday
Wednesday
day


Actual Productivity
sweing
finishing
department
finishing


Actual Productivity
Quarter1
Quarter2
Quarter3
Quarter4
Quarter5
quarter