

Homework4

November 20, 2018

1 Homework 4

Name: Yuqi Cao

1.1 Part1

1. Since a type I error is to falsely infer the existence of something that is not there, in this case would be finding the person pregnant when the person is really not pregnant. While a type II error is to falsely infer the absence of something that is. In this case would be finding the person not pregnant when the person is actually pregnant.

3.a) I would choose t-test. Since I want to evaluate how much different those two groups are.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- b) According to t-test's computation In this case, Prof. Rodriguez's experiment's numerator is greater than Prof. Seuss's result. But the denominator could be much greater than Prof. Seuss's. Therefore, after computation, t score of Prof. Rodriguez's is smaller than Prof. Seuss's.

1.2 Part2

4.

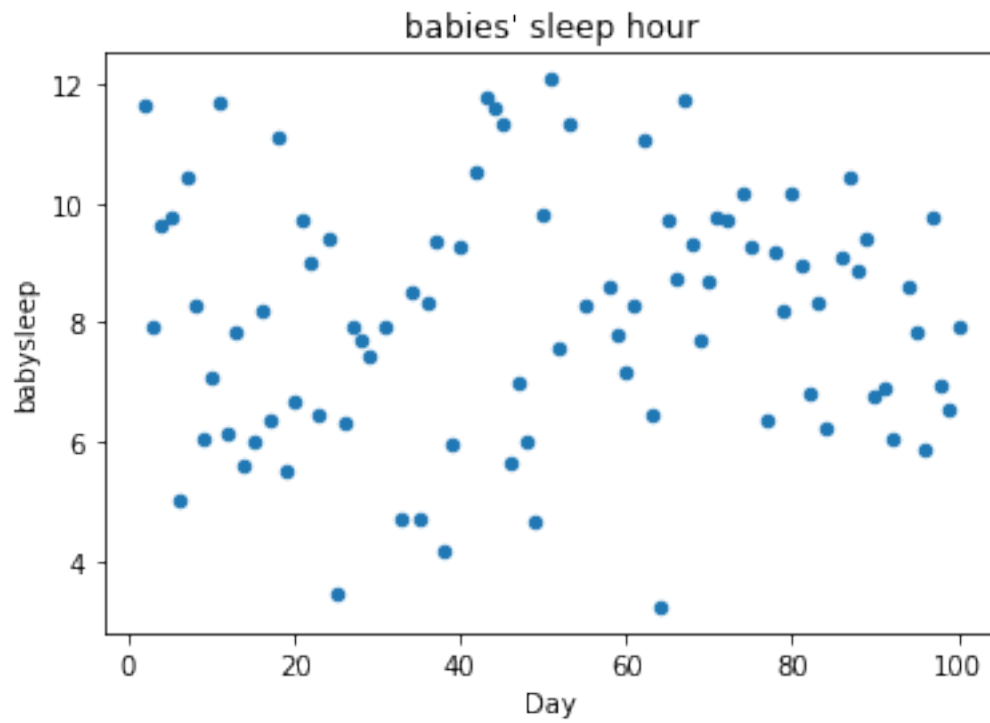
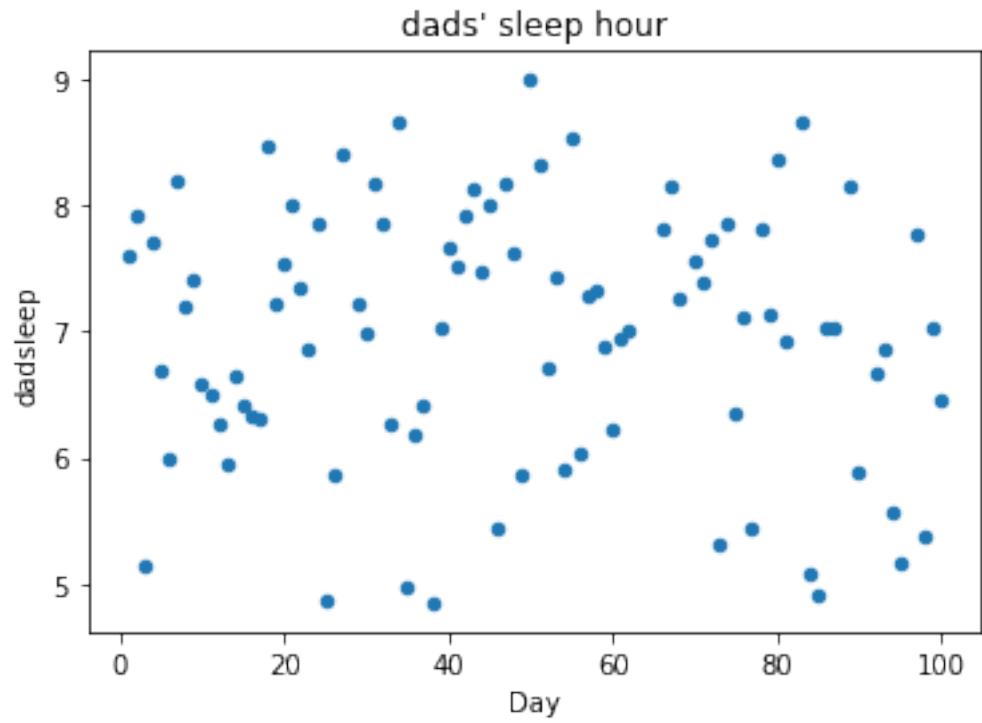
```
In [37]: import pandas as pd
import matplotlib.pyplot as plt

from IPython.display import HTML, display

df = pd.read_csv('parenthood2.csv', header=0,
names=['No.', 'dadsleep', 'babysleep', 'dadgrump', 'Day'])

# print(df)

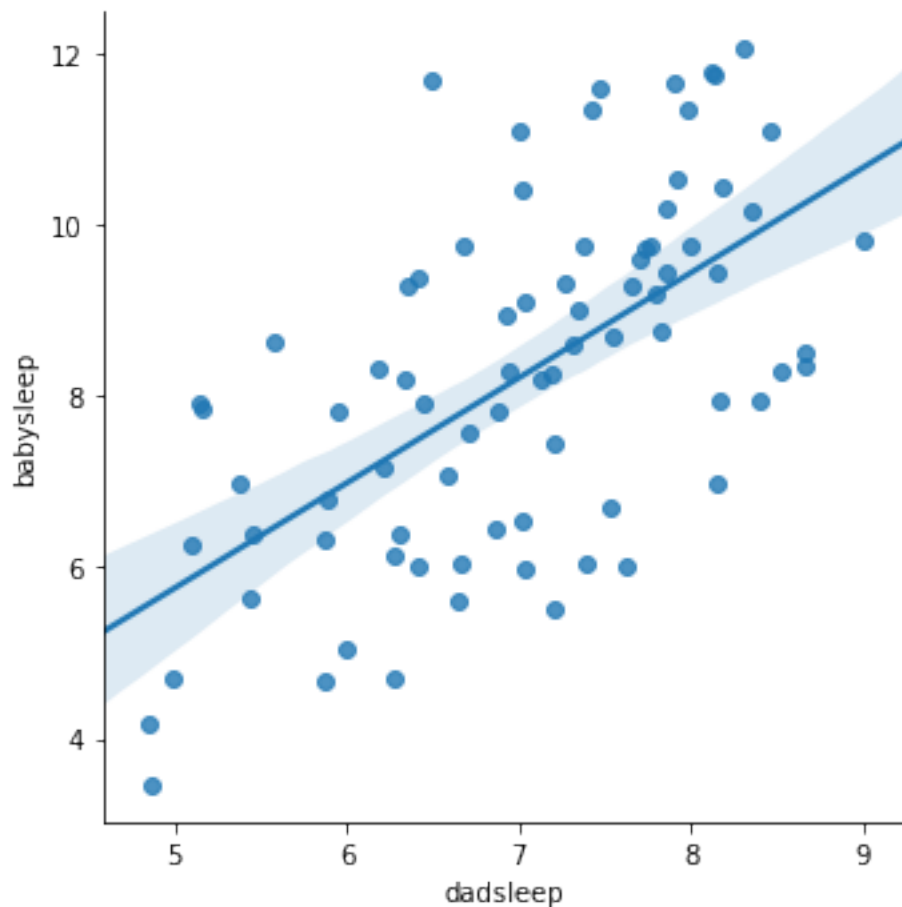
In [17]: import numpy as np
# i = np.linspace(0, 9, 18)
plot1 = df.plot(kind='scatter', x='Day', y='dadsleep', title='dads\' sleep hour')
plot2 = df.plot(kind='scatter', x='Day', y='babysleep', title='babies\' sleep hour')
```



2.the correlation coefficient between fathers' sleep and babies' sleep would be 0.615

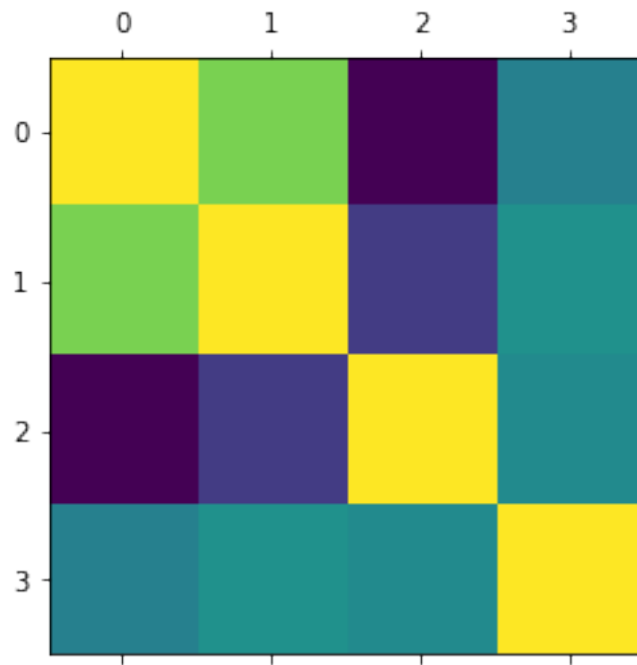
```
In [29]: import statsmodels.api as sm
import seaborn as sns
sns.lmplot(x='dadsleep', y='babysleep', data=df, fit_reg=True)
```

```
Out[29]: <seaborn.axisgrid.FacetGrid at 0xb7de9b0>
```



```
In [44]: # plt.matshow(df.corr())
newC=df[['dadsleep', 'babysleep', 'dadgrump', 'Day']].corr()
plt.matshow(newC)
print(newC)
```

| | dadsleep | babysleep | dadgrump | Day |
|-----------|-----------|-----------|-----------|-----------|
| dadsleep | 1.000000 | 0.614723 | -0.903442 | -0.076797 |
| babysleep | 0.614723 | 1.000000 | -0.567803 | 0.058309 |
| dadgrump | -0.903442 | -0.567803 | 1.000000 | 0.005833 |
| Day | -0.076797 | 0.058309 | 0.005833 | 1.000000 |



the strongest correlation of the pair would be dad's sleep hour and dad's grumpiness: - 0.903442