Exploring Neural Data final project

Introduction

During normal sleep the brain activity change between different states called REM, NREM (1-4) Movement time, and wake state. In this project we would like to investigate the effect of sleep-deprived night on the sleeping state during the night.
For the project a data set of four persons was used, the dataset contain EEG recording of two nights for each person, a well-rested night as baseline and recovery night that follows a sleep-deprived night.
The investigation method was to sum the total amount of each sleep stage in each night for each person and compare the two night of each person in order to identify similar patterns.

Study Questions

1. How the total sleep time affected from sleep-deprived night?
   Is there a pattern that the total sleep period in recovery nights (night after sleep-deprived night) is longer or shorter than sleep period in regular night (baseline)?
2. How much time it will take a person to fall asleep after sleep-deprived night?
   Does the time to fall asleep in recovery night is similar to the time it will take in a normal night?
3. How does each sleeping step type is affected from sleep-deprived night?
   Does the percent of each sleep step type (REM, NREM (1-4) Movement time, and wake state) during a recovery nights is kept the same as in regular night?

Dataset

The EEG recording is span over eight files, for each person a number is assigned (1-4), for each person there are 2 files, well-rested night (BSL) and recovery night (REC). for example file "S1_BSL.npz" relate to baseline of person one.
All files are in same structure.

File structure

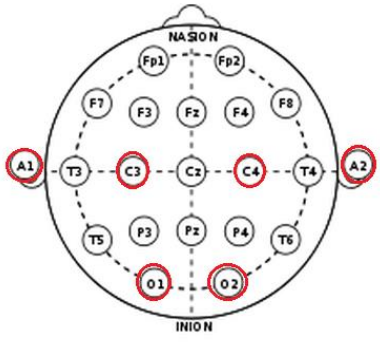
The file consists of three elements:

- "srate" – EEG sampling rate per second.
- "stages" – array of researchers classified information of the whole night stages. Each value in the array represents a 30 second state.
  - 0 - Awake
  - 1 - NREM Stage 1
  - 2 - NREM Stage 2
  - 3 - NREM Stage 3
  - 4 - NREM Stage 4
  - 5 - REM Sleep
  - 6 - Movement Time
  - 7 - Unscored (typically before lights out or after lights on, not analyzed)

- "DATA" – array of size nine containing the actual EEG in nine channels. Each channel is an array representing an EEG recording of different channel of the whole night.

Channels details

- EEG Channels - (See http://en.wikipedia.org/wiki/10-20_system_(EEG) for more information on the electrode naming conventions and for information about electrode locations.
  - Channel 1: C3/A2
  - Channel 2: O2/A1
  - Channel 8: C4/A1
  - Channel 9: O1/A2



- EOG Channels - ROC and LOC are the right and left outer canthi electrodes. The canthus is the corner of the eye.
  - Channel 3: ROC/A2 - right outer canthi electrodes
  - Channel 4: LOC/A1 - left outer canthi electrodes
- EMG Channels:
  - Channel 5: Chin EMG 1
  - Channel 6: Chin EMG 2
  - Channel 7: Chin EMG 3

Method

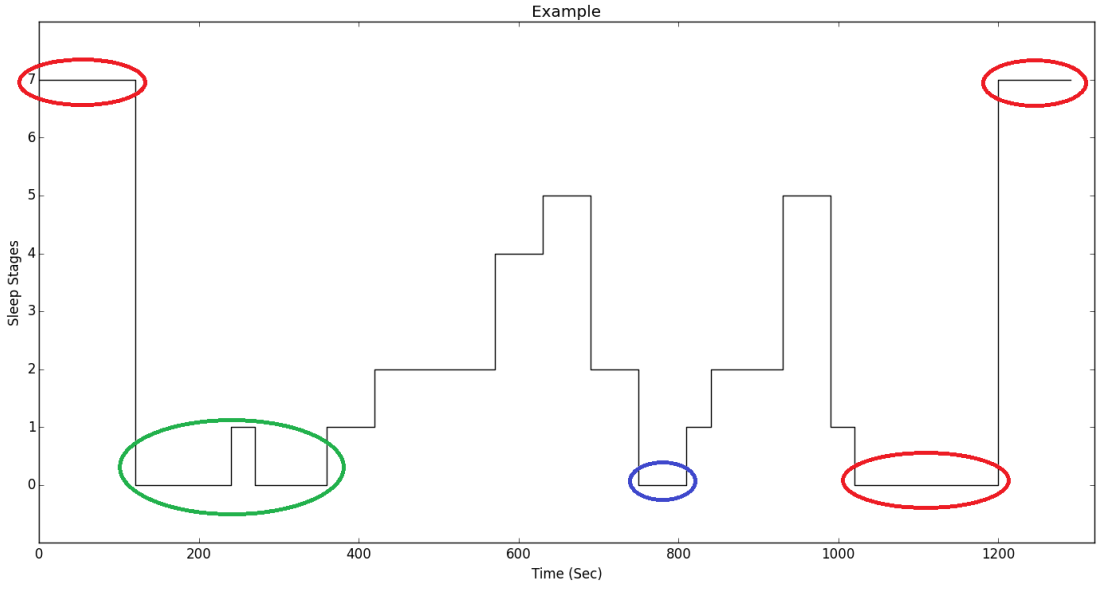
To answer both study questions I had used the pre-delivered researcher classified stages (see problems for reasoning).

In order to review only the sleep period I had created a cleansed EEG classification.
I had truncate the unclassified stages (7) from both start and end of EEG classification and the awakens period at the end of the EEG classification.
In addition I had separated the awakens period into two new stages (**Awake before sleep** and **Awake during sleep**).

- **Awake before sleep** – stages 0 and 1 from the beginning of truncated EEG classification until the first different stage found (2-6)
- **Awake during sleep** – any other awake stage (0) in the truncated EEG classification

Example of cleansed EEG classification creation

- **Red parts** are the part that truncated
- **Green part** mark as "Awake before sleep"
- **Blue part** mark as "Awake during sleep"



In order to answer question 1 I had compared the total sleep time of same person of baseline vs. recovery night. Total sleep time is the total time in cleansed EEG classification, except "awake before sleep" period.

In order to answer question 2 I had compared the "Awake before sleep" time of recovery night vs. baseline of the same person, in order to find a pattern for all persons.

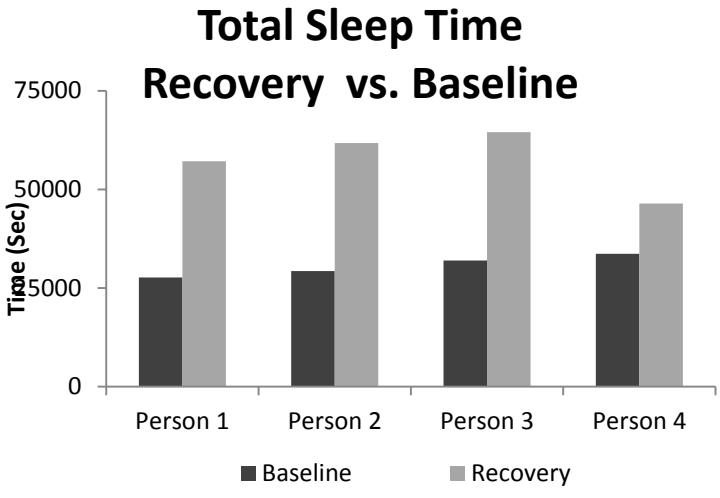
In order to answer question 3 I had sum each stage in each night and normalize it (dividing it with the total of cleansed EEG classification of that person & night). This gives the stage percent of that night. Then I had compared for same person each stage trying to find similar patterns for all persons.

Analysis

Question 1

How the total sleep time affected from sleep-deprived night?

Summaries data



	Baseline	Recovery
Person 1	31830	57300
Person 2	32670	61920
Person 3	33990	64710
Person 4	34830	46590

Reviewing the above data gives us a thick hint that recovery sleep time is longer than normal sleep time (as we probably would guess).

Hypothesis

H_0: $\mu_{REC} - \mu_{BSL} = 0$

H_A: $\mu_{REC} - \mu_{BSL} > 0$

This is a one side hypothesis with small number of elements. For that we will use the T distribution.

$n = 4$

$\overline{X}_{REC} = 57{,}630$ Sec

$\overline{X}_{BSL} = 33{,}330$ Sec

$S_{REC} = 6{,}901$

$S_{BSL} = 1{,}159$

$df = min(4 - 1, 4 - 1) = 3$

$(\overline{X}_{REC} - \overline{X}_{BSL}) = 57{,}630 - 33{,}330 = 24{,}300$

$$SE = \sqrt{\frac{{S_{REC}}^2}{4} + \frac{{S_{BSL}}^2}{4}} = 3{,}499$$

$$T_{df} = \frac{obs - null}{SE} = \frac{24{,}300 - 0}{3{,}499} = 6.945$$

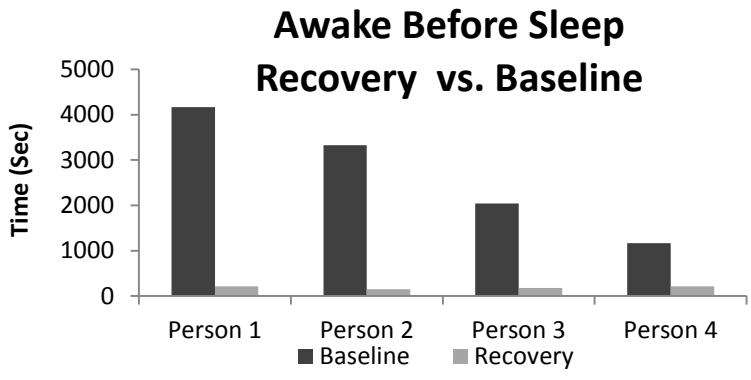
P-Value ≈ 0.003

Because of the low P-Value we would reject the null hypothesis and conduct that recovery sleep time (REC) is longer than normal sleep time (BSL).

Question 2

How much time it will take a person to fall asleep after sleep-deprived night?

Summaries data



	Baseline	Recovery
Person 1	4170	210
Person 2	3330	150
Person 3	2040	180
Person 4	1170	210

Reviewing the above data gives us a thick hint that in recovery nights the person fall asleep almost immediately (about 3 minutes), while in normal nights it takes a while to fall asleep (20 – 70 minutes).

Hypothesis

$H_0: \mu_{BSL} - \mu_{REC} = 0$

$H_A: \mu_{BSL} - \mu_{REC} > 0$

This is a one side hypothesis with small number of elements. For that we will use the T distribution.

$n = 4$

$\overline{X}_{REC} = 187.5\ Sec$

$\overline{X}_{BSL} = 2677.5\ Sec$

$S_{REC} = 24.9$

$S_{BSL} = 1{,}154.6$

$df = min(4 - 1, 4 - 1) = 3$

$(\overline{X}_{BSL} - \overline{X}_{REC}) = 2677.5 - 187.5 = 2490$

$$SE = \sqrt{\frac{{S_{REC}}^2}{4} + \frac{{S_{BSL}}^2}{4}} = 577.4$$

$$T_{df} = \frac{obs - null}{SE} = \frac{2490 - 0}{577.4} = 4.312$$

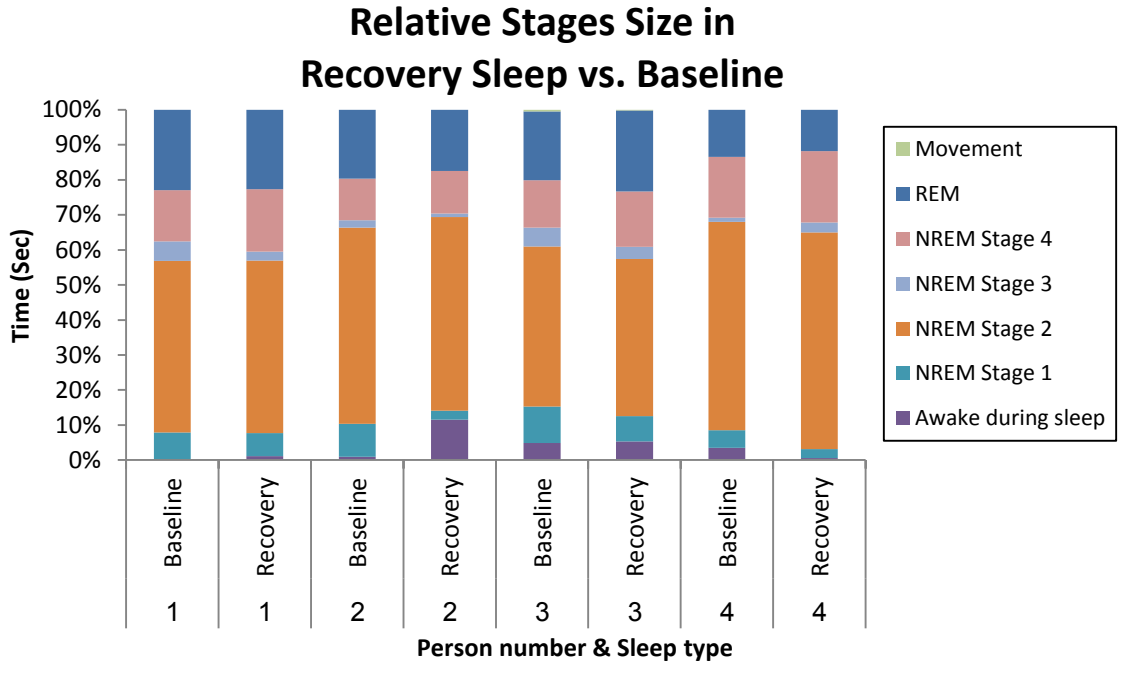
P-Value ≈ 0.011

Because of the low P-Value we would reject the null hypothesis and conduct that in recovery nights (REC) the person fall asleep faster than in normal sleep time (BSL).

Question 3

How does each sleeping step type is affected from sleep-deprived night?

Summaries data



Person number	Type	Awake during sleep	NREM Stage 1	NREM Stage 2	NREM Stage 3	NREM Stage 4	REM	Movement
1	Baseline	0.0	7.9	48.9	5.7	14.6	22.9	0.0
1	Recovery	1.1	6.6	49.2	2.5	17.9	22.7	0.0
2	Baseline	0.9	9.5	56.0	2.1	11.8	19.7	0.0
2	Recovery	11.6	2.6	55.2	1.0	12.1	17.5	0.0
3	Baseline	4.9	10.4	45.7	5.4	13.5	19.7	0.4
3	Recovery	5.3	7.2	44.8	3.5	15.8	23.1	0.2
4	Baseline	3.5	5.0	59.5	1.1	17.4	13.4	0.0
4	Recovery	0.6	2.6	61.9	2.8	20.3	11.8	0.0

Hypothesis

For each stage type a two side hypothesis test will be define. Since there are small number of elements, T distribution will use the.

$H_0: \mu_{BSL} - \mu_{REC} = 0$

$H_A: \mu_{BSL} - \mu_{REC} \neq 0$

$n = 4$

$df = min(4 - 1, 4 - 1) = 3$

Awake during sleep hypothesis test

$\overline{X}_{REC}$ = 4.647

$\overline{X}_{BSL}$ = 2.326

S_{REC} = 4.406

S_{BSL} = 1.952

$(\overline{X}_{REC} - \overline{X}_{BSL})$ = 4.647- 2.326= 2.321

$SE = \sqrt{\frac{{S_{REC}}^2}{4} + \frac{{S_{BSL}}^2}{4}}$ = 2.4

$$T_{df} = \frac{obs - null}{SE} = \frac{2.321 - 0}{2.4} = 0.9633$$

P-Value ≈ 0.4

Fail to reject the null hypnosis.

NREM Stage 1 hypothesis test

$\overline{X}_{REC}$ = 4.756

$\overline{X}_{BSL}$ = 8.196

S_{REC} = 2.194

S_{BSL} = 2.052

$(\overline{X}_{REC} - \overline{X}_{BSL})$ = 4.756- 8.196 = -3.44

$SE = \sqrt{\frac{{S_{REC}}^2}{4} + \frac{{S_{BSL}}^2}{4}}$ = 1.502

$$T_{df} = \frac{obs - null}{SE} = \frac{-3.44 - 0}{2.4} = -2.29$$

P-Value ≈ 0.1

Fail to reject the null hypnosis.

NREM Stage 2 hypothesis test

$\overline{X}_{REC}$ = 52.79

$\overline{X}_{BSL}$ = 52.54

S_{REC} = 6.417

S_{BSL} = 5.487

$(\overline{X}_{REC} - \overline{X}_{BSL})$ = 52.79- 52.54 = 0.248

$SE = \sqrt{\frac{{S_{REC}}^2}{4} + \frac{{S_{BSL}}^2}{4}}$ = 4.22

$$T_{df} = \frac{obs - null}{SE} = \frac{-3.44 - 0}{2.4} = 0.059$$

P-Value ≈ 0.95

Fail to reject the null hypnosis.

NREM Stage 3 hypothesis test

$\overline{X}_{REC}$ = 2.46

$\overline{X}_{BSL}$ = 3.568

S_{REC} = 0.902

S_{BSL} = 1.985

$(\overline{X}_{REC} - \overline{X}_{BSL})$ = 2.46 - 3.568 = -1.108

$SE = \sqrt{\frac{S_{REC}{}^2}{4} + \frac{S_{BSL}{}^2}{4}}$ = 1.090

$$T_{df} = \frac{obs - null}{SE} = \frac{-1.108 - 0}{1.090} = -1.016$$

P-Value ≈ 0.38

Fail to reject the null hypnosis.

NREM Stage 4 hypothesis test

$\overline{X}_{REC}$ = 18.53

$\overline{X}_{BSL}$ = 14.34

S_{REC} = 3.015

S_{BSL} = 2.022

$(\overline{X}_{REC} - \overline{X}_{BSL})$ = 18.53 - 14.34 = 2.192

$SE = \sqrt{\frac{S_{REC}{}^2}{4} + \frac{S_{BSL}{}^2}{4}}$ = 1.815

$$T_{df} = \frac{obs - null}{SE} = \frac{2.192 - 0}{1.815} = 1.208$$

P-Value ≈ 0.31

Fail to reject the null hypnosis.

REM hypothesis test

$\overline{X}_{REC}$ = 18.77

$\overline{X}_{BSL}$ = 18.92

S_{REC} = 4.602

S_{BSL} = 3.429

$(\overline{X}_{REC} - \overline{X}_{BSL})$ = 18. 77 - 18.92 = -0.148

$SE = \sqrt{\frac{S_{REC}{}^2}{4} + \frac{S_{BSL}{}^2}{4}}$ = 2.87

$$T_{df} = \frac{obs - null}{SE} = \frac{-0.148 - 0}{2.87} = -0.0516$$

P-Value ≈ 0.96

Fail to reject the null hypnosis.

Sleep movement hypothesis test

$\overline{X}_{REC}$ = 0.046

$\overline{X}_{BSL}$ = 0.11

S_{REC} = 0.08

S_{BSL} = 0.191

$(\overline{X}_{REC} - \overline{X}_{BSL}) = 0.046 - 0.11 = -0.064$

$SE = \sqrt{\frac{S_{REC}{}^2}{4} + \frac{S_{BSL}{}^2}{4}} = 0.104$

$$T_{df} = \frac{obs - null}{SE} = \frac{-0.064 - 0}{0.104} = -0.618$$

P-Value ≈ 0.58

Fail to reject the null hypnosis.

Summary of hypothesis tests results of question 3

From the above hypothesis tests it seems that there is no pattern of change in sleep step type percent during the night sleep between recovery nights and normal night

Summary Results

From the above dataset we come to the following conclusions

1. Recovery sleep night is longer than normal (baseline) sleep night time.
2. In a recovery sleep night person fall asleep almost immediately while in normal (baseline) sleep night it may take a while until a person will fall asleep.
3. There are no evidence for a significant effect of sleep-deprived night over the percentage of sleep step types of a full night sleep.

Problems

I started this project by building my own classifier and reach about 70% matches to the researcher classified stages, since it is relatively low and the study result may be unreliable, I sadly decided to use the pre-delivered researcher classified stages.
Having better classifier can enable me in follow-up analysis to investigate additional datasets which do not have pre-delivered classification.

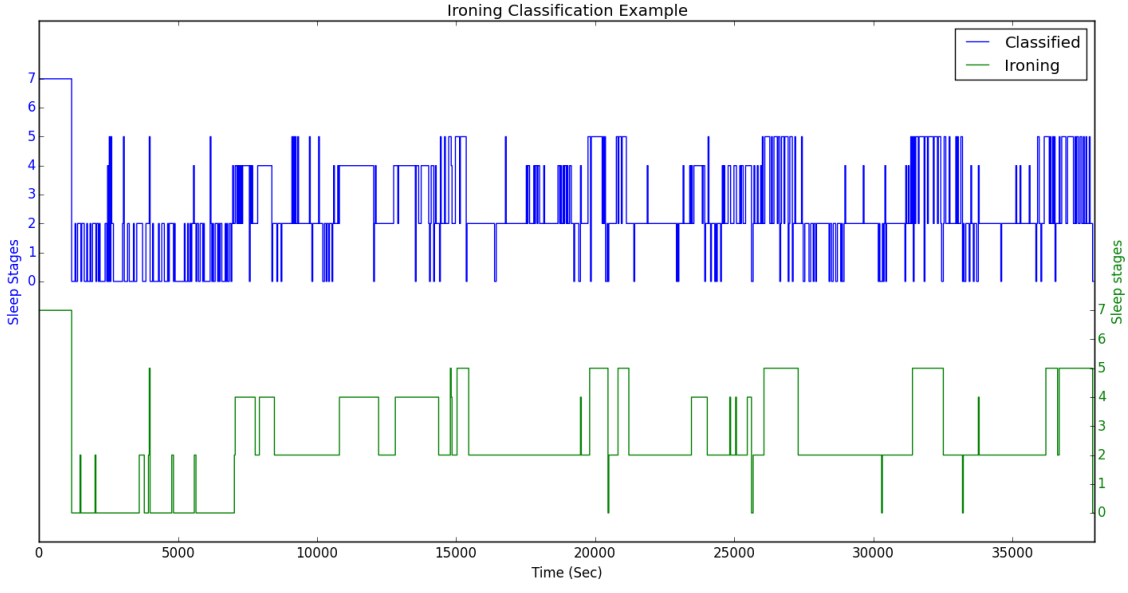
Follow-up Analysis

Since the number of sample is low (4 samples of recovery and baseline), the obvious next step would be to extend the number of samples and redo the same analyses.

Other major step would be to extend my classifier in order to achieve better accuracy. This will enable uses of additional unclassified EEG recording.

Interesting programming tricks

Since my classifier was relatively inaccurate it creates strong "waves" when plotting the classification (see example below). In order to overcome this issue I develop method called "iron" which flattens those waves. In the example below you can see the original classification in blue, and the "Ironed" classification in green.



The rationale was that the stages tends to be study for a while and then change.

The algorithm for this method is as follows:

For each level starting from the less populated level to the most populated do the following:

Starting from the beginning of the classified array scan the values

If current value different from the input level continue to next.

Otherwise loop while in level or end of array.

If the current value is not in level (could be spike or true point exiting this level), do the following:

Record the number of continuous non level and check if it had exceeded the limit (typically more than 3).

If yes verify that the majority of the array subset has value equal to level, if this is the case set this subset to level.

Code sample:

```
def  iron(classified,level,sequenceMaxBreak=3):
"""
Flatten the classified array, erasing local small picks (less than sequenceMaxBreak)
and creating continues levels.
"""
        i=0
        classifiedLen = len(classified)
        while i < classifiedLen-1:
                # if current value not equal to value continue
                if classified[i] != level:
                        i += 1
                else:
                        levelStart = i
                        inLevel = True
                        outOfLevel = 0
                        # while in level and not end of array
                        while inLevel and i < classifiedLen-1:
                                # Progress one step
                                i += 1
                                # if current value equal to value start count from zero continues out of level
                                if classified[i] == level:
                                        outOfLevel = 0
                                # Otherwise increase continues out of level
                                else:
                                        outOfLevel += 1

                                # if continues out of level excide the limit than get out of level (exit loop)
                                if outOfLevel > sequenceMaxBreak:
                                        inLevel = False

                        # count number of values in this subarray
                        isValue = 0
                        for j in range(levelStart,i):
                                if classified[j] == level:
                                        isValue += 1

                        # if majority equal to level set to level
                        if isValue > (i-levelStart)/2:
                                for j in range(levelStart,i):
                                        classified[j] = level
                        # Otherwise do the same as above starting from the next point
                        # after the first level value in sequence
                        else:
                                i = levelStart + 1

        # return the "iron" classified array
        return classified
```