# Group Project Journal for Sprint3

Name: Peng Ye/ Chen Shen/ Anna
Team: No. 18.
The Master of Sprint3: Chen Shen
Date: 12.04.2018 –- 20.04.2018

## 1. Project Management

Our group's sprint3 started from 12.04.2018. Actual start date is one week later than originally planned because of the holiday in April. The sprint3 mainly includes data analysis, testing and CSS designing. In sprint3, we have the daily meeting at 9 pm every day and the extra review meeting that helped us review the whole project is held on 19.04.2018.

### 1.1 Design, Planning Notes and Materials

1.1.1 The plan of data analytic features:

TASK1a: Automatically getting data from Mysql database (once / week)

TASK1b: Train data and Train the Machine Learning Algorithm with one mode

TASK1c: Make the prediction on the bike usage condition, the input data could be weather (Rain/None Rain, temperature, humidity, Time duration, e tc.)

TASK1d: Illustrate the prediction data onto webpage

1.1.2 The plan of testing features:

TASK2a: Test the python code by the unit test

TASK2b: Test the display of web application

1.1.3 The plan of CSS designing:

TASK3a: Use simple CSS layout to landscape our webpage

### 1.2 Scrum project management (Trello board, slack logs, google docs, Etc)

Our group mainly use slack and face-to-face communication to manage our project. First, we will do coding in our local machines for each small task and then update it to the Github. Finally, we will download the code from Github to EC2 and make it run in the EC2. After finishing one small task we will do this job and see whether the task has been finished. And we will start next task when we use slack to tell all that we have finished the last task.

### 1.3 Meeting notes (log of daily standup, notes from sprint reviews, retrospectives)

1.3.1 Sprint3 meeting notes:

Date: 12 April 2018

Place: O'Briens Science Centre at 9:30

Present: Dr.A.Lawlor, P. Wicke Chen Shen, Peng Ye, Anna Ryzova

Q1: Update all the features and backlogs and in the meeting we need to talk about what the rest features are.( Like CSS and Testing or any others)

Note: We should add the CSS to our web pages. And Peng has already upgrade the testing feature and data analysis feature into small tasks.

Action: Chen and Peng

Q2: Do we need to add a reviewing meeting in the end of the project?

Note: 4.19 From 9:30

Review all the work. And invite Phil to join our meeting.

Q3: What features or tasks we need to do in the sprint3? (Allocation of Tasks)

Notes: In the next week, we should work together to finish the data analysis part and the CSS part which includes the code integration. In order to work together, we try to arrange time to sit around the table every day and do the coding job in the same version

Action: Chen, Peng and Anna

Q4: Finally, what we need to submit to the teacher? (Both for group and individuals)

Postpone to the next week's meeting.

Action: None

1.3.2 The reviewing meeting notes

Date: 19 April 2018

Place: O'Briens Science Centre at 9:30

Present: Dr.A.Lawlor, P. Wicke Chen Shen, Peng Ye, Anna Ryzova

Q1: The review of the whole project.

Note: We talked about the problems in the processing of this project. And we reviewed the our work in sprint3.

Q2: How to write and organize the documents?

Note: Follow the instructions in comp30670-2018-group.pdf and the Data Storage.pdf.

1.3.3 Daily standup:

**Date          Peng/Chen/Anna**

| 4.12 | Peng |
|---|---|
| | Today: |
| | 1.    Have Sprint 2 review meeting and Sprint3 meeting |
| | 2.    Update backlog combing with Sprint1 part and Sprint3 part |
| | 3.    Wrote Scrum meeting notes |
| | Tomorrow, |
| | Work on 4 tasks of Data Analytics Part |
| | Chen: |
| | Today, we have the sprint 2 review meeting and sprint 3 meeting. And i am writing the Sprint3 meeting notes. But i am not very sure about whether should we continue our plans? |
| | Tomorrow, we will work together on data analysis from 10am in science building |
| | Anna: |
| | I am working on chart 4 today.   Tomorrow we will discuss how to work together on the Data Analytics part tasks. |
| 4.13 | Peng |
| | I will keep on working Data analytic tonight and tomorrow |
| | Chen: |
| | Today, i helped Peng with the display of the prediction page. And i finish the sprint3 meeting note and update the Backlog by adding the Intergration features. |
| | Tomorrow, we will continue to work on data analysis and try to finish this part. And Peng and me will meet in the science building at 10 am tomorrow. |
| | Anna: |
| | Today we had an extra meeting on our final project stage.   We allocated the tasks for the Data Analytics feature.   I start on Task 26 from our backlog today. |
| 4.14 | Peng: |
| | Today, |
| | Chen and I worked together on Data Analyzing of Bike Station part. We finished task 24 in backlog, and 80% of   task 25 as well. In addition, we finished integration of web page and I have update bike5.html onto github: |
| | http://mytest:5000/bike5/ |
| | I will try to finish task 25 and task 26 in backlog, maybe include some testing works as well |
| | Chen: |
| | Today, |
| | 1. I worked with Peng on the data analysis. And i trained a model of station1 and predicted the usage of the bikes by the 5 |

weather features.

2. We also have integrated the earlier codes into one version web application.

Tomorrow,

1. i will continue to integrated the codes and separate the js codes to the .js files. And i will also spend some time on CSS or heat map.

Anna:

I was doing my Task 26 today and about to commit the code that will read the JSON code from Dublin_chart_7 and display it.

Tomorrow will continue.

| | |
|---|---|
| **4.15** | Peng: |

Today,

1. I finished all tasks on data analytic part. You guys can have a test with this page:

http://ec2-34-217-16-180.us-west-2.compute.amazonaws.com:5000/bikeReview/

2. worked with Chen on CSS, *,js files's integration on EC2 instance

tomorrow, I will work on some testing work and I will start on document's collection from next Monday on

Chen:

Today,

1. i separate the javescript code to the .js files and update the github

2. i almost finished the heatmap code

Tomorrow,

i will work on the CSS features with you guys

Anna:

Today I was testing the files.

| | |
|---|---|
| **4.16** | Peng: |

Today, I didn't do any coding works. tomorrow, I think we may have a quick review on what we have done and what we missed over last week.

Chen:

Today,

1. i did nothing for our project.

Tomorrow,

I think we need to talk about how to write the document together and finish the tasks of CSS feature.

Anna:

I will continue with Chart 4 tonight.   We will decide tomorrow what tasks have to be completed in our project and allocate them.

| | |
|---|---|
| **4.17** | Peng: |

Today, I changed code on data analytics part according to philip's suggestions:

1)     Use python to dump training model onto particular files

2)     Adjust the layout of data analytics   page

Tomorrow, I will work on document and CSS

Chen:

Today, i have spent two hours on CSS and our webpage's main colour is blue and white

Tomorrow, we will start writing documents and finish all the code tasks.

Anna:

I spent some time today preparing my Learning Journal 3, notes and all documentation I have up to date.   I will upload it to Github until Friday

| | |
|---|---|
| **4.18** | Peng: |

Today's work:

1.     Worked with Chen on CSS/Layout integration on flask server. Modified and added new codes and completed the final bits

on Bike Station webpage. (5 hours in total)

2.      Layout testing work is finished. Only a few text issues to be adjusted.

Tomorrow:

1.      Have discussion with Philip on document's writing

2.      Unit testing code writing and integrating to master bit

Chen:

Today, i worked with Peng to finish the CSS of our wep app

Tomorrow, i will finish the project journal of sprint3 and start writing the personal journal.

Anna:

Hi guys, I will attend tomorrow practical

## 1.4 Feature selection/product backlog for sprint3

| Sprint 3 | | | |
|---|---|---|---|
| Data Analytic front-end layout 1: Submitting and features selection | P1 | Peng/Chen | Completed |
| MySQL Data query from both tables, need to get the following values: Humidity, Temp, weather, wind-speed, station number, Bank-stands, available bike stands, last update.MySQL Data query from both tables, need to get the following values: Humidity, Temp, weather, wind-speed, station number, Bank-stands, available bike stands, last update. | P1 | Peng/Chen | Completed |
| Model training and particular predications on each station, based on the following continuous and category features: temperature, humidity, weather, wind-speed, and clock (1,2,3,4,5…24) | P1 | Peng/Chen | Completed |
| Data Analytic front-end layout 2: illustration with chart on particular station's bike usage (such as how many bikes are available in the particular bike station in different time clock everyday, based on the given conditions) | P1 | Anna | Completed |
| Data front-end layout 3: add new button linking between the first google map and the second google map. | P1 | Peng | Completed |
| Verify back-end can get the necessary data from weather API timely (12 hours/ time) and insert them into MySQL DB | P3 | Peng/Chen | Completed |
| Verify back-end can get the necessary data from Bike Station API timely (5 minutes/time) and insert them into MySQL DB | P3 | Peng/Chen | Completed |
| Verify the weather are displaying current temperature exactly on the webpage | P3 | Anna | Completed |
| Verify the weather are displaying historical temperature exactly on the webpage | P3 | Anna | Completed |
| Verify chart1 and chart2 display the bike information clearly and exactly | P3 | Peng/Chen | Completed |
| Verify chart4 display the bike information clearly and exactly | P3 | Peng/Chen | Completed |
| Verify chart6 display the correlation between bike & weather clearly and exactly | P3 | Peng/Chen | Completed |
| Verify Data Analytic webpage can get the request values precisely | P3 | Peng/Chen | Completed |
| Verify Data Analytic can show the predicted bike information precisely and clearly | P3 | Peng/Chen | Completed |
| Verify the operating and setting out of Flask server is controllable | P3 | Peng/Chen | Completed |
| Verify the outlooking of webpage is good | P3 | Peng/Chen | Completed |
| Verify the log in back-end is accessible | P3 | Peng/Chen | Completed |
| Intergrate the code into one page for webpage | P1 | Chen/Peng | Completed |
| Change js quary into the .js file and link to index page | P2 | Chen | Completed |
| Add CSS layout into our page and also some pictures about bike | P2 | Chen/Peng | Completed |
| Write new unit test cases on Python code (dataAnalytic class) | P2 | Peng | Completed |

## 1.5 Burndown charts, (sprint and/or release) for sprint3
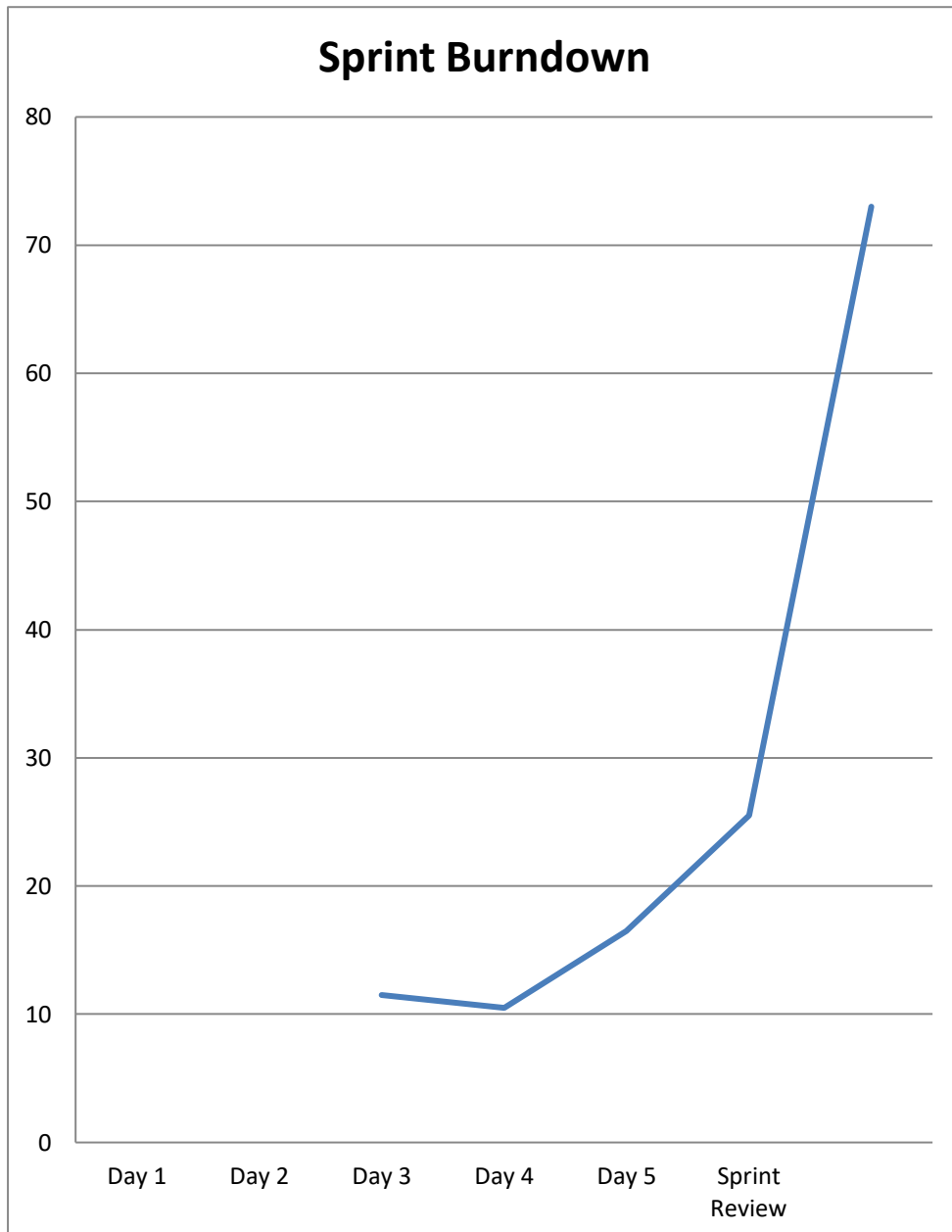
**1.6 Managing deliverables, prototypes**

1.6.1    Data analysis:

Bike Station 34 analytics on Rain from 2018-04-13 20:14:34 to 2018-04-02 09:31:35
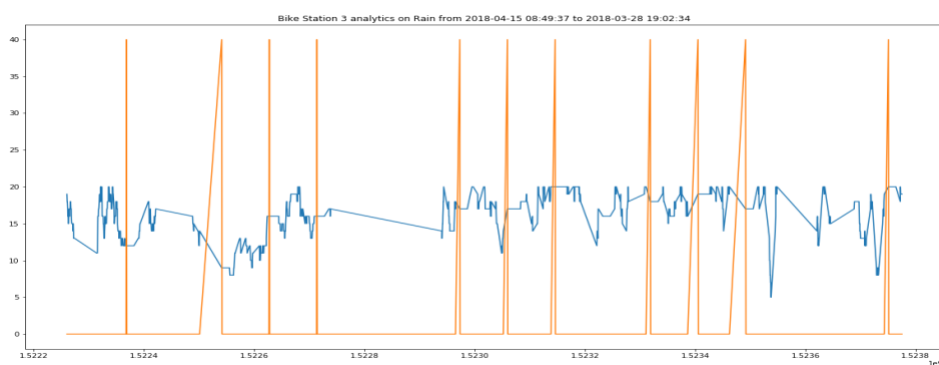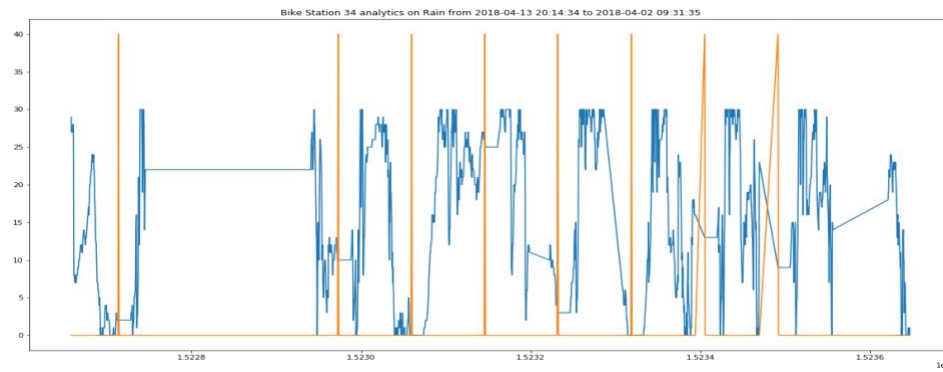
The two pictures draw the tendency of available bikes in station 3 and 34 when the weather is rain. And these pictures show that the model can be trained to predict the available bike tendency.

The deliverables:



**Prediction of Available Bikes in Dublin**

Temperature :(from -10 to 30)
Humidity:     (from 70 to 100)
WindSpeed:    (from 0 to 20)
Bike Station NO.:      (1-104)
SELECT A WEATHER ▲▼     Predict

*This is the prediction form, where you can provide the day's weather conditons and station number here. After clicking 'Predict' button, you will be able to get the predicted bike usage condition of that bike station in that day. You can also click the Markers which containing the station numbers on above google map, by which it will pop up the details of that Bike Station.*

The result of the prediction:



**Prediction of Available Bikes in Dublin**

Temperature :(from -10 to 30)
Humidity:     (from 70 to 100)
WindSpeed:    (from 0 to 20)
Bike Station NO.:      (1-104)
SELECT A WEATHER ▲▼     Predict

**Bike Staton NO.100('HEUSTON BRIDGE (SOUTH)') will be used 16/25 bikes in average on that day**

### 1.6.2  Unit testing for python code:



```
            if os.path.isfile(file):
                os.unlink(file)
            my = dataAnalytic.dataAnalytic()
            my.importModel(stationNumber)
            # Verify the dump file for station's training model has be created
            self.assertTrue(os.path.isfile(file))

        def test_dataAnalyticClass_003(self):
            '''Vefify getPredictionOnStation() method in dataAnalytic Class'''
            temperature=10
            humidity=0
            weather='Rain'
            windSpeed=5
            stationNumber=1
            my = dataAnalytic.dataAnalytic()
            prediction, bikeStands = my.getPredictionOnStation(temp=temperature,
                                                    humidity=humidity,
                                                    weather=weather,
                                                    windSpeed=windSpeed,
                                                    stationNumber=stationNumber)
            self.assertTrue(bikeStands >= prediction and prediction >= 0)
            self.assertTrue(isinstance(prediction, int) is True)
            self.assertTrue(isinstance(bikeStands, int) is True)

if __name__ == '__main__':
    suite = unittest.TestLoader().loadTestsFromTestCase(TestAPI)
    unittest.TextTestRunner(verbosity=2).run(suite)
ubuntu@ip-172-31-24-252:~/SEProject_Group18/tests$ ls
conftest.py   unit_test.py
ubuntu@ip-172-31-24-252:~/SEProject_Group18/tests$ vi unit_test.py
ubuntu@ip-172-31-24-252:~/SEProject_Group18/tests$ ls
conftest.py   unit_test.py
ubuntu@ip-172-31-24-252:~/SEProject_Group18/tests$ python unit_test.py
test_dataAnalyticClass_001 (__main__.TestAPI)
Verify getBikeStation() method in dataAnalytic Class ... ok
test_dataAnalyticClass_002 (__main__.TestAPI)
Verify importModel() method in dataAnalytic Class ... ok
test_dataAnalyticClass_003 (__main__.TestAPI)
Verify getPredictionOnStation() method in dataAnalytic Class ... ok

----------------------------------------------------------------------
Ran 3 tests in 3.077s

OK
ubuntu@ip-172-31-24-252:~/SEProject_Group18/tests$
```

1.6.3   Display testing:

This is the instance link and our web application is put in here:

http://ec2-34-217-16-180.us-west-2.compute.amazonaws.com:5000/

When click the link, we can see the following picture:
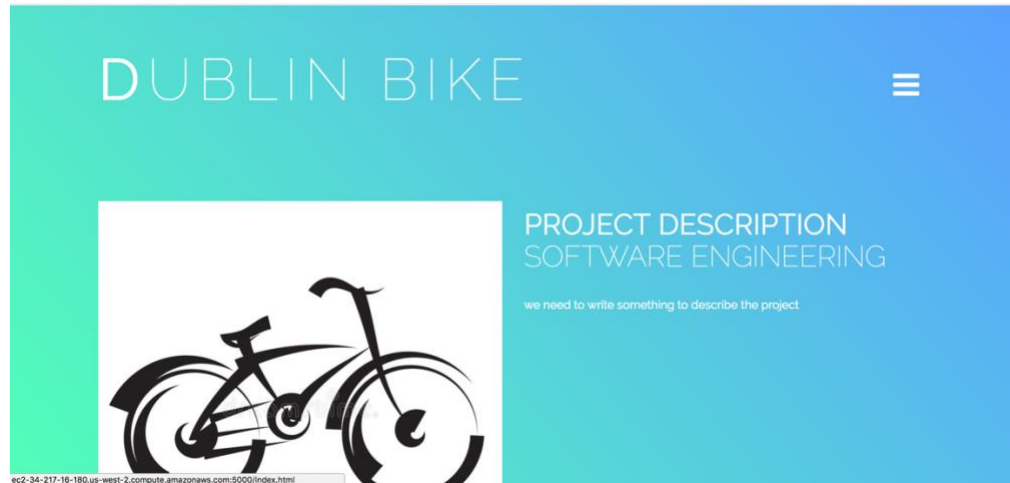


1.6.4   The model used in data analysis:

```
                          OLS Regression Results
==============================================================================
Dep. Variable:     Available_bike_stands   R-squared:                       0.441
Model:                            OLS   Adj. R-squared:                  0.435
Method:                 Least Squares   F-statistic:                     74.59
Date:                Sat, 14 Apr 2018   Prob (F-statistic):           1.68e-57
Time:                        17:07:07   Log-Likelihood:                -1631.6
No. Observations:                 478   AIC:                             3275.
Df Residuals:                     472   BIC:                             3300.
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                          coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept              -260.5048     22.703    -11.474      0.000    -305.117    -215.893
C(WeatherMain)[T.Rain]    6.4403      1.509      4.269      0.000       3.476       9.405
Temp                      6.7872      0.839      8.090      0.000       5.139       8.436
Humidity                  2.1522      0.178     12.116      0.000       1.803       2.501
WindSpeed                 1.3353      0.126     10.628      0.000       1.088       1.582
Rain3H                   -2.8000      1.221     -2.294      0.022      -5.199      -0.401
==============================================================================
Omnibus:                       16.816   Durbin-Watson:                   0.121
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               10.978
Skew:                          -0.235   Prob(JB):                      0.00413
Kurtosis:                       2.425   Cond. No.                     6.50e+03
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 6.5e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

## 2. Skills Practiced

In sprint3, we focus on the data analysis and testing. As a result, the following skills have been practiced:

2.1 Python coding skill

2.2 The process of data analysis

2.3 Handle the huge data frames

2.4 Unit testing

2.5 CSS coding

2.6 Writing project reports

## 3. Learning Done

The most important we have learned in sprint3 is cooperation. In the sprint1 and sprint2, we worked for the project individually. For example, one person worked on the whole feature like backend and other people did not care a little about this feature. And I think it is inefficient for the group to finish the project. Also, we cannot learn new knowledge because we all did the part we are good at. However, we, especially Peng and Chen, always worked together to solve the problem and we all participated all the parts in sprint3. And we find it is efficient for the project.

Another thing we have learned is that we can separate the big feature into several small tasks. Then, we can finish these small tasks step by step.

## 4. Goals

In the future assessment, we want to set the following goals:

1. Better communication with each other

2. Better management of project

3. Complete tasks on time

4. Improve the skill of coding