Team meeting minutes for Accelerating group

Date: Thursday, August 22, 2024

Before this meeting, we have reviewed 6 selected articles regarding the area of rental market.In this meeting, we discussed about our findings, the topic of our research project, how we are going to conduct this research, agenda of the research, tasks allocation and how this research can contribute to the client. Details are in the following.

1. Findings

We found that there are few studies in this area. Our aim is to apply data science methods to analyze Australian rental market. But, most of the studies of this kind focus on the rental market in China or America. And the studies regarding Australian rental market didn't use the data science methods. This clearly shows that, our research project is of originality.

We also found contents related to:

- how to collect data using Python
- how to clean data using Python
- how to use machine learning models to analysis dataset

Detail of these findings can be seen in:

https://github.com/users/DiaoFu/projects/1/views/1?pane=issue&itemId=75137743

2. Agenda of this research project

Based on the above findings, we decided that the **topic** of our research project is :

The determining factors of rental price in ACT (the area is restricted to ACT at the current stage).

In this process, we will be using CKAN to support our data management(including data collecting, data cleaning and data storage). Then we will generate 'CKAN analysis report' to reflect on our experience of using CKAN to support our research. We will comment on what are the benefits of using CKAN and what are the shortcomings that can be improved by fixing bugs or adding extensional function.

The Agenda of this research is as following:

1) Data collection

We will conduct web scrapping on the website Domain.com.au to collect rental data.

We expect this to be down **in w5 and w6.**

2) Data management

We will use CKAN to store and manage our data.

We expect this to be down **in w5 and w6.**

3) Data cleaning

We will use Python to clean our data. This process involves the using of CKAN to support.

We expect this to be done **in w6.**

4) Generating of CKAN analysis report 1
Based on the above experience of using CKAN, we will generate a first analysis report of CKAN. We will state our experience of using CKAN to manage our data, describe how CKAN supports our tasks of data collection, data cleaning, and then comment on the advantages of using CKAN. We will also provide advice on how CKAN can be improved.
We expect this to be done **in w6.**

5) Data analysis
We will use machine learning models to analyze the dataset collected.
We expect this to be done **in w7.**

6) Generating and Presenting of research report
We will use the results obtained above to generate a research report. Then we will present such research result using visualization tools and web designing techniques.
We expect this to be done **in w9.**

7) Generating of CKAN analysis report 2
After completing the above stages, we will generate the second and final analysis report of CKAN. We will state our experience of using CKAN to manage our data, describe how CKAN supports our tasks and then comment on the advantages of using CKAN. We will also provide advice on how CKAN can be improved.
We expect this to be done **in w10.**

3. How this research project will add value to our client
Firstly, as what has been described that, this project can provide advertising effect for our client and assist our client in marketing.

Secondly, this research will provide CKAN a user reflection regarding the advantages of using CKAN and shortcomings of CKAN. This can help CKAN to be reinforced by maintaining its advantages and improving its shortcomings.

4. Task allocation for the next weeks
1) Data collection:
This task is assigned to Yuxin Mu, Diao Fu, Anbo Wu.
Expected result: A CSV file containing the data needed. The data should include features like rental price, location, size of house, number of bed room.
Due date is 9 pm, Monday, August 26, 2024

2) CKAN installing and environment setting up
This task is assigned to Chuang Ma.
Expected result: CKAN is installed and is ready to use. Chuang Ma is able to upload data into CKAN.
Due date is 9 pm, Monday, August 26, 2024

3) Data management using CKAN
This task is assigned to Chuang Ma,Yuxin Mu, Diao Fu, Anbo Wu.
Expected result: Use CKAN to store and manage our data. Ensure data is consistent.
Due date is 9 pm, Tuesday, August 27, 2024

4) Data Cleaning
This task is assigned to Yuxin Mu, Songxuan Li, Doeun Han, Qifeng Zhang.
Expected result: a clean dataset without error and missing value.
Due date is 9 pm, Tuesday, August 27, 2024

5) Generating of CKAN analysis report 1
This task is assigned to Yuxin Mu, Anbo Wu, Songxuan Li, Doeun Han, Qifeng Zhang.
Expected result: provide the first analysis report for client
Due date is 9 pm,Wed, August 28, 2024

6) Data analysis
This task is assigned to Yuxin Mu, Songxuan Li, Doeun Han, Qifeng Zhang.
Expected result: Conduct data analysis on dataset generated
Estimated Due date is 9 pm, Tuesday, Sep 10, 2024