STA 490 Final Team Project Guidelines

Cheng Peng

West Chester University

Contents

Introduction

This is essentially a capstone team project that involves multi-item survey data, survey instrument validation, information aggregation of multi-item survey data, SAP drafting, EDA and regression modeling, statistical analysis reporting, and HTML presentation.

Data Sources

Three data sets are suggested for the team projects.

Employee Turnover Data

The data set and its description can be found at https://pengdsci.github.io/datasets/#employeeturnover.

The following five variables measure the personality from different perspectives. Justify whether it makes sense to combine these variables to define a new and practically meaningful (combined) index variable.

- **extraversion** - (outgoing/energetic vs. solitary/reserved)
- **independence** - Agreeableness (friendly/compassionate vs. critical/rational)
- **selfcontrol** - Conscientiousness (efficient/organized vs. extravagant/careless)
- **anxiety** - Neuroticism (sensitive/nervous vs. resilient/confident)

- **novator** - openness to experience (inventive/curious vs. consistent/cautious)

```
EmployeeTurnOver = read.csv("https://pengdsci.github.io/datasets/EmployeeTurnOver/turnover-data-set.csv
```

Employee Attrition Data

The data set and its description can be found at https://pengdsci.github.io/datasets/#employeeattrition.

The following five variables measure employee satisfaction from different perspectives.

- **EnvironmentSatisfaction**
- **JobSatisfaction**
- **RelationshipSatisfaction**
- **WorkLifeBalance**

```
EmployeeAttrition = read.csv("https://pengdsci.github.io/datasets/EmployeeAttrition/EmployeeAttritionDa
```

Project Ideas

The above two data sets are related to employee retention. The following are the potential questions you may consider when analyzing these data sets.

Employee Attrition Data

Practical questions related to the **Employee Attrition Data**.

General Questions for EDA

- Any missing value issues? Need imputations? justification?

- Is there any potential issue of multi-collinearity among predictor variables?

- Is it possible to aggregate the information in these satisfaction-related variables to create a new *combined satisfaction variable*.

- Explore inter-relationship between satisfaction variables.

Primary Questions

Identify intervenable factors that are associated with

- employee attrition,
- engagement, and
- performance.

Secondary Questions

Identify potential intervenable factors that impact the level of employee satisfaction from different perspectives.

- Job satisfaction
- Environmental satisfaction
- Relationship satisfaction
- Work-life balance satisfaction

Other Findings

Document all practically important findings that were not asked to address any of the clients' initial questions and make relevant recommendations.

Employee Turnover Data

Practical questions related to the **Employee Turnover Data**

General Questions for EDA

- Any missing value issues? Need imputations? justification?

- Is there any potential issue of multi-collinearity among predictor variables?

- Is it possible /necessary to aggregate the information in personality-related variables to create a new *combined personality index variable*. Justify your answer.

Primary Questions

Identify intervenable factors that are associated with

- employee turnover

Final Capstone Project Guidelines

The steps and requirements for the capstone projects are outlined in the following.

Project Teams (based on WCU student ID number)

- **Team One**: Angelo, Ian, Jack;
- **Team Two**: Jaiden, Samantha, Gianna;
- **Team Three**: Kyle, Haley, Tyree;
- **Team Four**: Alice, Joshua G, Joshua Z.

Project Topics

Three topics are associated with **employee attrition data**. To be more specific, you are expected to identify intervenable factors that are associated with/predict

1. employee attrition,

2. engagement, and

3. performance.

One is associated with **employee turnover data**. To be more specific, you are expected to identify intervenable factors that are associated with/predict

4. employee turnover.

Team Assignment

There will be NO individual capstone project. Each team will work on one of the above four projects.

Time Lines

1. **SAP - due 4/28 (week 13)**: Each team should prepare a statistical analysis plan (SAP) and submit it to D2L at the end of the week of April 22. We will not meet in class on April 25. Members of individual teams are expected to work together to prepare the SAP. Individual team members' contributions must be clearly outlined at the end of the report.

2. **Team HTML Presentation Document - due 5/5 (week 14)**: Prepare your team report based on the SAP and an HTML presentation and publish them on your GitHub (Each team member should submit copies of the team report and the presentation). <font color = "red">No class meetings will be scheduled this week.

3. **Formal Oral Presentation - Thursday, 5/9 from 10:30 AM - 12:30 PM**: Each team will have 30 minutes to present the analysis. The major areas to be assessed are statistical rigor, effectiveness of presentation, aesthetics of the presentation design, and individuals' engagement in QA sessions.

Rubrics

Two rubrics of project report and presentation can be found on the course web page.