Project Proposal - Social Mobility

Perach Hen Elkayam

5.4.2022

Motivation

This project deals with the issue of social mobility in Israel. **Social mobility** is the movement of individuals, families, households, or other categories of people within or between social strata in a society. It is a change in social status relative to one's current social location within a given society. Because of the structure of the population in our country that includes a big variety of cultures, I believe that is easier to isolate variables that might indicate a possible improvement of social status. I find this issue very interesting because I can understand what it is needed from someone in order to improve and change his social status. In this project I would like to find answers to the questions:

- Can we predict future social status by the background of the participant and his parents.

- Trying to characterize similar groups by features of social mobility.

- Can we identify variables that influence social mobility.

Description of the data set

The data is a survey conducted by *The Central Bureau of Statistics in 2018. The study participants are 7,500 people representing the population in Israel aged 20 and over. I got the raw data set from the Central Bureau of Statistics.*

The data consist of 402 questions in the subject of Social Mobility. In order to give description on the data set I divided the questions to 12 groups:

1. <u>Details about household members</u> - questions about marital status, country of birth, age, sex, etc.

2. <u>Details about apartment, car and Housekeeping</u> - These questions focus on property, apartment ownership (and a number of owned apartments), ownership of a private vehicle (and a number of vehicles per capita in the household). These questions also ask if is there an assistant or caregiver at home and for what purposes.

3. <u>Positions on residence</u> - Description of attitudes regarding the living area, such as satisfaction with public transportation services, street security, air pollution, garbage collection, satisfaction from the relationship with neighbors, etc.

4. <u>Health and lifestyle</u> - Questions about the health background of the study participant, chronic diseases, disabilities, and whether there is difficulty in daily functioning.

5. <u>Religion and ethnic origin</u> - Religious self-definition, religious background and questions concerning ethnic origin.

6. Skills: Studies, languages, courses, military service and driver's license - Questions about academic studies and professional training, number of years of schooling, education among the parents of the research participant and whether he did military or national service.

7. Computer and Internet use and access to technology - Questions about using mobile phones and computers. Use of the Internet and for what purposes.

8. Employment - Employment status, type and the field of employment, scope of work, job satisfaction, professional future, accessibility to workplaces, impact of change in family status, medical status or place of residence on employment, etc.

9. Vacation, contact with family and friends and volunteering - Questions about social life, going out for fun and vacations, satisfaction and confidence in social relationships and contact with family members. In addition, questions about volunteering and contributing to the community.

10. Economic status and income and welfare- Income, economic conduct, perception regarding the economic situation of the survey participant from a young age to the present day.

11. Civic involvement - Involvement in political and public life, participant opinion on civic bodies such as the health system, the education system, the Israel Police, etc.

12. Trust in government institutions - Participant opinion on government institutions such as government, the justice system, and the health system. Does the participant encounter discrimination in various areas of life.

    Score for social status - There are different definitions for social status and they may vary between different cultures. In modern terms and in western societies measurable characteristics for social status are education and economic status (income and property). I would like to make a new variable "social status score" that combines the two in order to make a statistical analysis.

Research Questions as Statistical Learning Problems:

- **Supervised Learning:** Prediction of social status score (combination of education and economic status) by the status of the participant and his parents on the past (for example, education and levels of religiosity).

- **Unsupervised Learning:** Can we find similarities between different participants by indications of their social status and split them to small amount of groups? If so, it might help us understand better the relationship between different characteristics of social status.

- **Inference:** Can we estimate the effect of a specific variable (such as education) over social mobility by controlling the confounders (trying to find participants with similar backgrounds and compare those with and without education).

Methods:

- **Supervised Learning:**

  - Logistic Regression: I can investigate the coefficients easily.
  - Tree method: I can draw the decision tree easily.

- **Unsupervised Learning:**

  - KNN: Easy to understand, I can control the number of groups.
  - PCA: In order to investigate dimension reduction.

- **Inference:**

  - Logistic Regression with stratification over different populations in order to control the confounders.

Evaluation and Experimenting:

- **Supervised Learning:**

  - I will evaluate the performance of the algorithms by familiar performance metrics for prediction accuracy and simplicity of the model (hopefully the models will accurate and easy to interpret). I'll be using cross-validation methods for hyper parameter tuning for the tree model and stepwise selection for logistic regression.

- **Unsupervised Learning:**

  - I will evaluate the performance of the algorithm by trying to find consistency between my results and some known knowledge in the field of social mobility.

- **Inference:**

  - If the variable of interest will still be significant after controlling for the confounders by stratification of the population I will asses that the effect of the variable is in fact consistent with the theory and not just correlated with other significant variables.