

R-Bootcamp Report

Bernardo Freire Barboza da Cruz & Leonid Gavriluk

03 February, 2022

Contents

1	Introduction	2
2	Extract and Transform	3
2.1	Data Import	3
2.2	Data Exploration & Visualization	4
2.3	Merging Data	4
2.4	Data Visualization	4
2.5	Fit Model	4
2.6	Chapter of Choice TBD	4
3	References	4

1 Introduction

Describe the use case. . . .

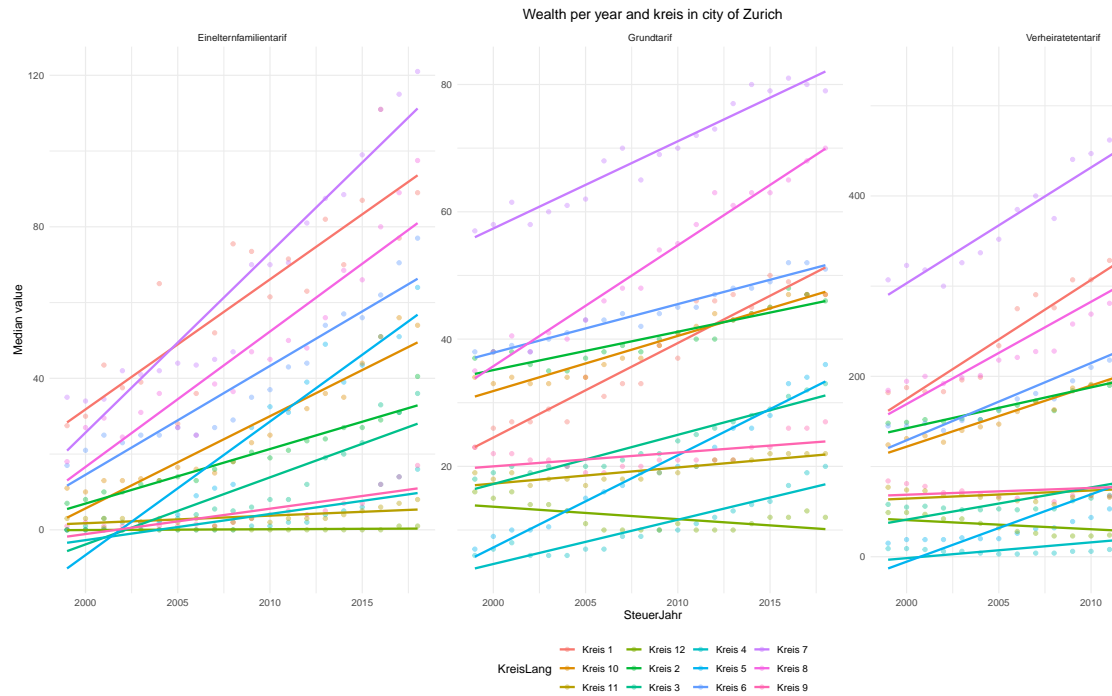
1. Most active districts
2. Party dynamics 2006-2018
3. Activity of genders (e.g. women are active voters)
4. Activity of age groups
5. Relationships between education and activity.
6. Relationships between income and activity.

2 Extract and Transform

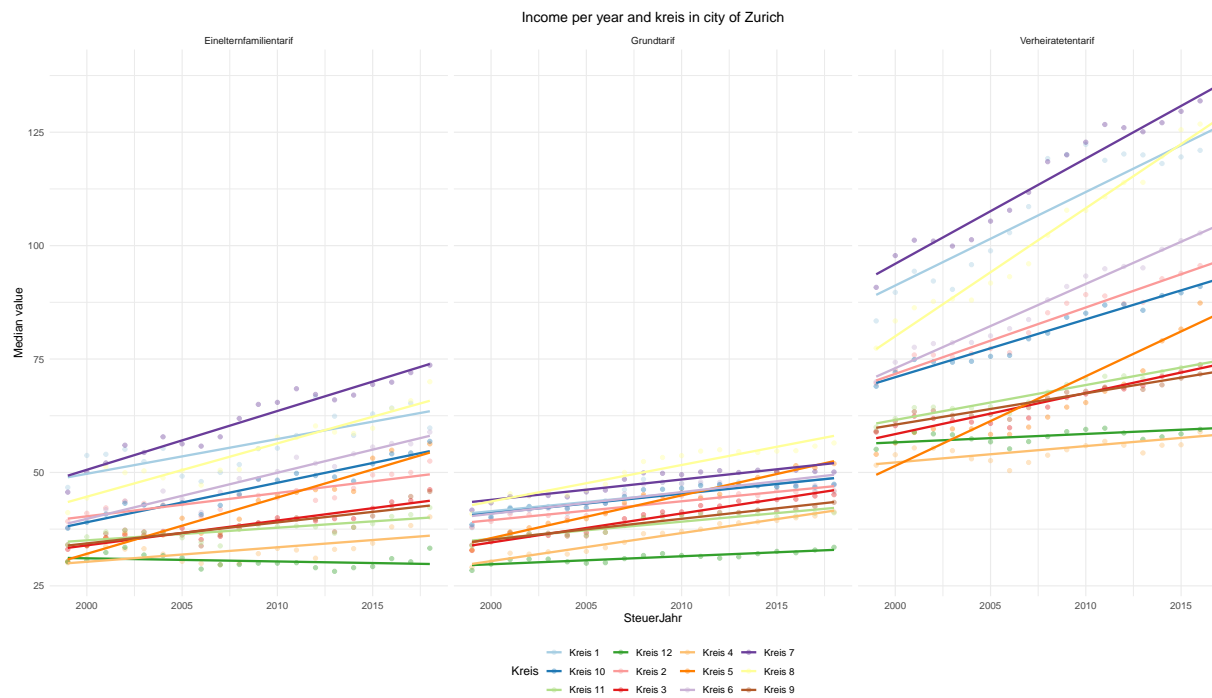
2.1 Data Import

The data is imported from *Stadt Zürich Open Data*.

2.1.1 Exploration: Wealth and Income data



2.1.1.1 Wealth Data



2.1.1.2 Income

2.1.2 Exploration: Education

2.1.2.1 Preprocess Data The encoding of the column `RaumSort` consists of the Kreis and further local details. The last digit describes neighborhood of each Kreis. It means:

- Two digits: The first digit describes the Kreis and the second the neighborhood
- Three digits: The first *two* digits describes the Kreis and the third the neighborhood

Add a new column to dataset for Kreis and neighborhood.

2.2 Data Exploration & Visualization

2.3 Merging Data

2.4 Data Visualization

2.5 Fit Model

2.6 Chapter of Choice TBD

3 References