Predicting Local Economic Activity based on Historical Employment Data

Hamed Asasi

Capstone Project, Data Incubator

Goal and the Users

- Goals:
  - An interactive visualization tool illustrating historical business activity across the U.S.
  - Determine which industries best predict economic activity
  - Short-time (2-3 years) prediction of other economic measures: income, poverty, housing prices

- Users:
  - Local governments
  - Individuals planning to move into the area or start a small business

Datasets

- **Main**: census data on *Statistics of U.S. Businesses (SUSB)* :
  - Employment and payroll data for businesses in over 350 metropolitan statistical areas (MSA) over the years 1997-2017
  - Information on size of businesses
  - Type of industry (construction, manufacturing, finance, etc.)
  - Over a million data values

- Census data on *Small Area Income and Poverty Estimates (SAIPE)*
  - Over 3000 counties
  - 1995-2019

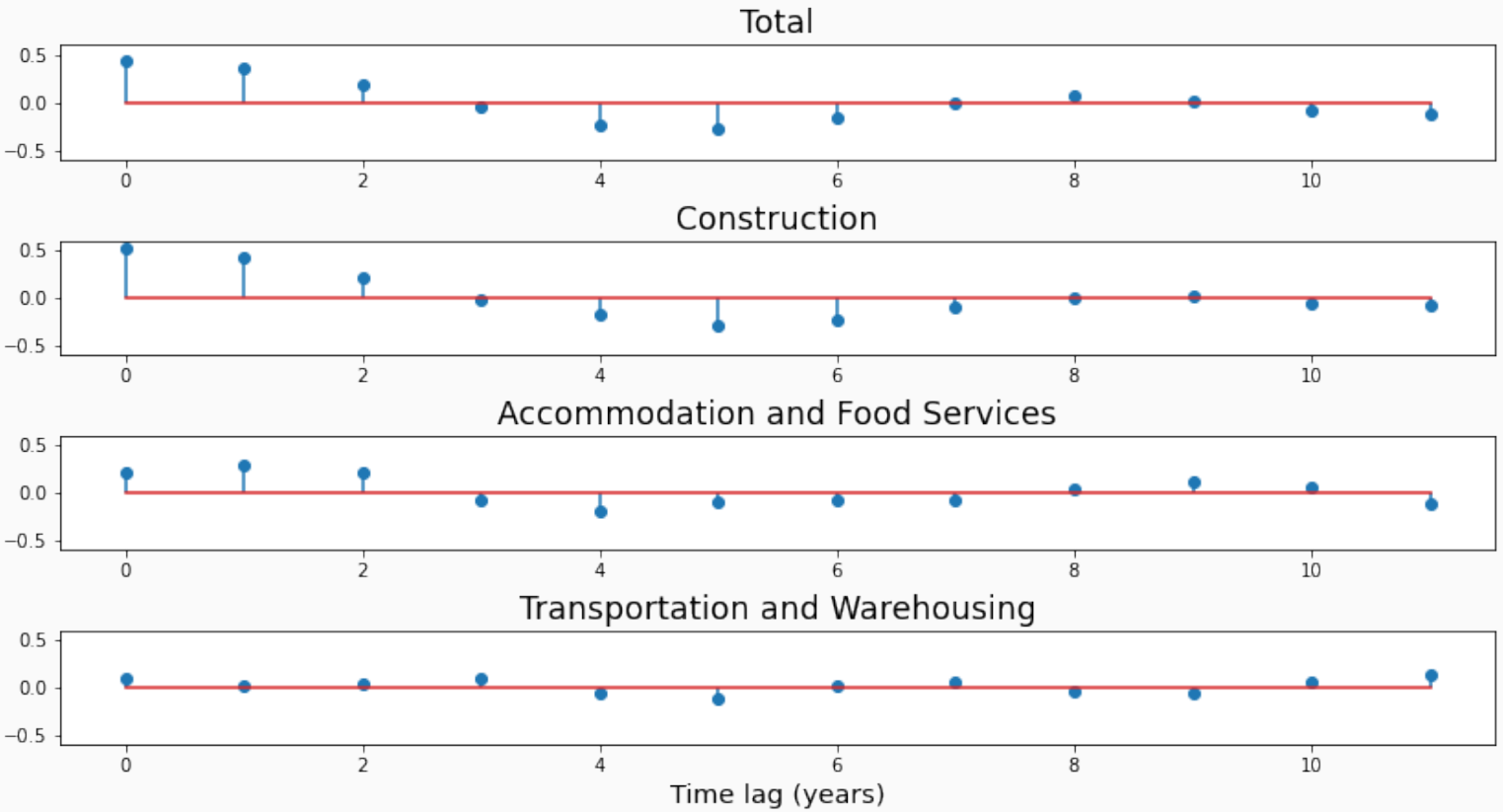
- Zillow's home value index

Challenges

- Time series but with few time data
- Extract features:
  - Separate employment data by industry, size of the business, population of MSA
  - Isolate the local trends from state/country trends
- Match metropolitan areas to states and counties

Exploratory Analysis

- Correlation of Employment vs. housing prices
  - Point percentage change relative to the country



Plan

- Separate time-local correlations from seasonal variations
- Build a prediction model with
  - Features: aggregation of employment data for each industry at a specific time
  - Target values: percentage change in employment, housing, poverty and median income for the next few years
- Train: linear regression, regression with random forests.

Thank You!