Final Project 1: Lightning Presentation

DS-SF-30

Ivan Ledyashov

Efficiency of Wine Market – Problem

- Value of wine: quality vs. price
- Market assumes it is efficient: price a predictor of quality?
  - Price ↑ quality ↓ - poor value
  - Price ↓ quality ↓ - poor/average value
  - Price ↑ quality ↑ - average value
  - Price ↓ quality ↑ - high value
- What is quality? Are other variables a better predictor of value?

Efficiency of Wine Market – Data

Several free APIs and static datasets are available:

- api.wine.com / api.snooth.com – wine retailer / site for oenophiles
  - Price Min, Price Max, Year, Appellation, Varietal, Wine Type, Vineyard, Reviews, Geo Location, Ratings
- archive.ics.uci.edu/ml/datasets/Wine - Institute of Pharmaceutical and Food Analysis and Technologies
  - Fixed Acidity, Volatile Acidity, Citric Acid, Residual Sugar, Chlorides, Free Sulfur Dioxide, Total Sulfur Dioxide, Density, pH, Sulphates, Alcohol, Malic Acid, Ash, Alcalinity of Ash, Magnesium, Total Phenols, Flavanoids, Nonflavanoid phenols, Proanthocyanins, Color Intensity, Hue, OD280/OD315 of Diluted Wines, Proline, Quality

Efficiency of Wine Market – Hypothesis

- Price is not perfectly correlated with quality
- Quality, price and value may be predicted by year, varietal, appelation, vineyard, geo location, wine descriptors, volatile acidity, pH, alcohol or the combination of these variables

Resources:

- https://onlinecourses.science.psu.edu/stat857/node/223
- http://archive.ics.uci.edu/ml/datasets/Wine
- http://archive.ics.uci.edu/ml/datasets/Wine+Quality
- http://rstudio-pubs-static.s3.amazonaws.com/175762_83cf2d7b322c4c63bf9ba2487b79e77e.html
- https://github.com/allanbreyes/udacity-data-science/blob/master/p3/submission.Rmd
- http://caseyiannone.com/Wine_Exploratory_Analysis/

Running Pace Adjustments for Elevation Gain – Problem

- Are you running up and down hills as fast as you can?
- Running up hills requires more energy, causing muscles to use more oxygen, straining lungs and producing more lactic acid than can be metabolized by the body
- Running down hills puts more strain on knees and joints, and requires runner to use a set of different muscles to control stride and cadence
- Pace increase constraints and enablers are: heart rate, resting HR - max HR, cadence, stride length, elevation grade, weather, etc.

Pace Adjustments for Elevation Gain – Data

- Runkeeper – GPS (GPX) files exportable by time range
  - Time, latitude, longitude, heart rate, steps, pace, elevation at every GPS point
- Strava – API access
- Fitbit – activity data in CSV and GPS data in TCX export
- Garmin – activity data in CSV and GPS data in TCX export

Pace Adjustments for Elevation Gain – Hypothesis

- Running pace adjusts more when running up hill than when running downhill - uphill speed may be decreased and downhill speed increased without sacrificing pace
  - Elevation ↑ running pace ↓
    - Heart rate ↑ (increases less during cold weather)
      - Cadence ↓
  - Elevation ↓ running pace ↑
    - Heart rate ↓ (decreases less during cold weather)
      - Cadence ↑

Leading Indicators for Credit Rating Change – Problem

- Company's changing financials and unexpected events are the primary reason for credit rating adjustments
- Which of the fundamentals are a better predictor credit rating adjustment?
  - Price to Book Value
  - Enterprise Value to Free Cash Flow (EV/FCFF)
  - Price to Earnings (P/E)
  - Revenue Growth
  - Operating Cash Flow Growth
  - Debt to Equity
- Which of non-fundamentals are also a predictor of rating adjustment:
  - Employee Salaries
  - Number of Employees
  - Stock Price
  - Stock Beta

Leading Indicators for Credit Rating Change – Data

- Intrinio – financial statement and market data, API
  - https://intrinio.com/
- SEC – financial statement data from EDGAR going back to 2009: XBRL, zipped by quarter
  - https://www.sec.gov/dera/data/financial-statement-data-sets.html
- Standard & Poor's – credit ratings: XBRL, zipped XMLs per company for the last 15 years
  - https://www.standardandpoors.com/en_AU/web/guest/regulatory/ratingshistory

Leading Indicators for Credit Rating Change – Hypothesis

Non-financial indicators are a predictor of credit rating adjustment

- **Requirements:** You should present from a prepared deck that explains the following:
  - ***The Problem***: What's the background and scope of the project idea? What problem are you attempting to address or solve? Who may it matter to?
  - ***The Data***: What data exists to help solve this problem? Where is it coming from? What does the data look like? What is the observation?
  - ***Your Hypotheses***: Given the problem and data you're aware of, what do you believe is the solution? What does success look like?
- **Detailed Breakdown:** Your presentation must:
  - Explain three potential projects, including: a problem statement, a hypothesis, and potential datasets.
  - Be two to three minutes in duration.
  - Demonstrate familiarity with the domain of the data.
- **Bonus:**
  - High-level data points can help show the "improvement" scope that's available.
  - Cite similar/previous work that will better inform your strategy that can be taken.