Data Preparation

Filtering Out

- negative quantity
- negative price
- december 2011 data

Data Imputation

- null description
- null customer id

Time Period Mapping

Stock Data

- Dec 2010 period 1
- Nov 2011 period 12

New Column

revenue = quantity *price

Customer Data

• unique customer + period in each row

- unique stock + period in each row
- transactions with multiple same customer + period or stock
- + period aggregated: average (quantity, price, revenue) and total (quantity, revenue)



Customer Data Predictions

Stock Data Predictions

- predicted average (quantity, price, revenue) and total (quantity, revenue) for each unique customer + period or stock + period for the following time periods: 13 (December 2011), 14 (January 2012), 15 (February 2012)
- predictions based off customer ID, country, and Time Period for customer data
- predictions based off stock code, count-vectorized description, and Time Period for stock data