

Lab 2 - Linguistic Survey

Stat 215A

Jonathan Fischer

October 7, 2014

1 Introduction

2 Redwood Data

2.1 Temperature Density Estimates via Kernel Smoothing

2.2 Loess Smoothed Humidity vs Temperature

3 Linguistic Data

3.1 Data quality and cleaning

This dataset isn't as bad as the redwood data, but there are still some issues. You should discuss them here and describe your strategies for dealing with them.

3.2 Exploratory Data Analysis

This is where you compare pairs of questions with discussion and plots.

4 Dimension reduction methods

This is where you discuss and show plots about the results of whatever dimension reduction techniques you tried - PCA, hierarchical clustering, K-means, random projections, etc.

5 Stability of findings to perturbation

6 Conclusion



