

# Statistics for Machine Learning and Big Data

An Introduction

Welcome

Jeff Abrahamson

7 April 2015

# Welcome

# Useful facts

I am a computer scientist.

# Useful facts

Python is a useful tool.

Our support SLA





## **ADVISORY FOR VISITORS TO THE PARK**

Because of the many natural hazards in Hawaii Volcanoes National Park, visitors should stay alert and pay attention to any warnings that might be provided to them on entering the park, at the Visitor Center, by warning signs throughout the park, by park personnel, or by sirens in special areas of the park. Hazards include, but are not limited to, volcanic eruptions, lava flows, earthquakes, tsunamis, poisonous gases, hot steam vents, rockslides, the sudden collapse of shoreline areas into the ocean, sharp rocks, dangerous footing on some hiking trails, and surfaces that may collapse on contact to uncertain depths and possibly into molten lava in some areas. Because of the extent of the park and the unpredictable nature of volcanic and seismic activity, warnings of these hazards may not always be available. Please do not disturb, and be cautious around, any monitoring instruments, cables, and wires that might be encountered while hiking or camping in the park.



**Those hiking or camping in shoreline areas should be especially aware of tsunami hazards.** Should you hear a siren in a coastal area move as quickly and safely as possible to higher elevations. You may only have seconds to save your life. Be alert for any unusual changes in sea level as a possible indication of a tsunami. You may not feel the ground shake if the source of the tsunami is far away. Large locally generated tsunamis have struck the coastal areas of the park in 1868 and 1975 with vertical heights on land of up to 47 feet above sea level. If possible, camp at higher elevations. If necessary, plan an evacuation route before sundown and be alert for any sirens, unusual sounds from the ocean, or other warnings. Remember that a tsunami is a series of waves and the first wave is not necessarily the largest. The time interval between waves can be from about 5 to 20 minutes. Note that for a number of reasons, functioning sirens may not exist in your hiking or camping area. Also, in some instances, functioning sirens may operate for only a few seconds before they are damaged by a tsunami.

# Course structure

## Today

- Probability and statistics (maybe mostly review)
- *Break*
- Distributions and inference
- *Lunch*
- Regression
- *Break*
- *Frequentists and Bayesians*

# Course structure

## Tomorrow

- Decisions and clustering
- *Break*
- Spam