



# Visualizing Data from a MOOC

Katelin Bailey    Jialin Li    Naveen Kr. Sharma

Computer Science & Engineering, University of Washington

## Problem

- Visualizing course data from a PL Coursera class.
- Compare statistics across two offerings.
- Provide insights into the data that are valuable to instructors



```
fun append (xs, ys) =
  if null xs
  then ys
  else (hd xs):: append (tl xs, ys)

fun map (f, xs) =
  case xs of
    [] => []
  | x :: xs' => (f x)::(map (f, xs'))

val a = map (increment, [4,8,12,16])
val b = map (hd, [[8,6],[7,5],[3,0,9]])
```

## Motivation

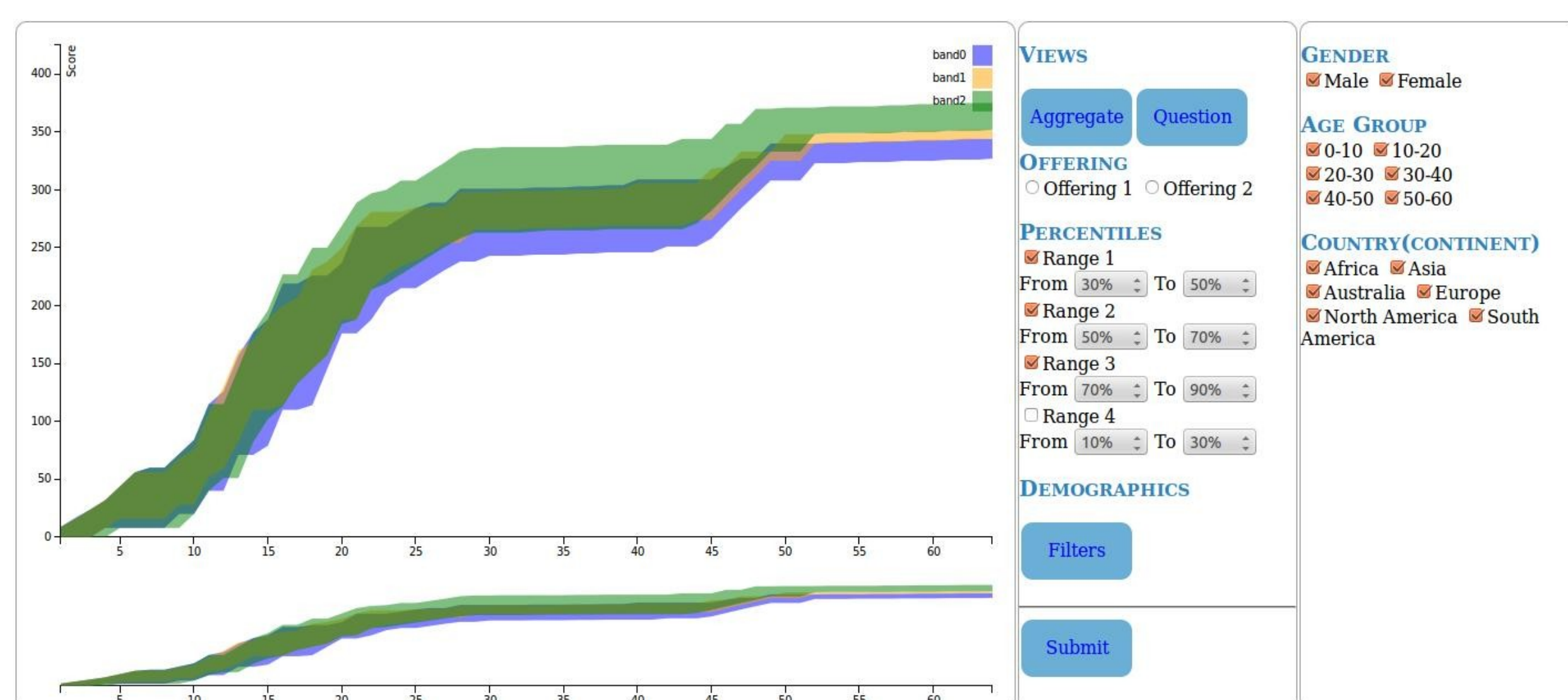
### Emerging trends

- Open on-line courses are becoming more popular.
- Enrollment number increases every year.

### Need to understand the data

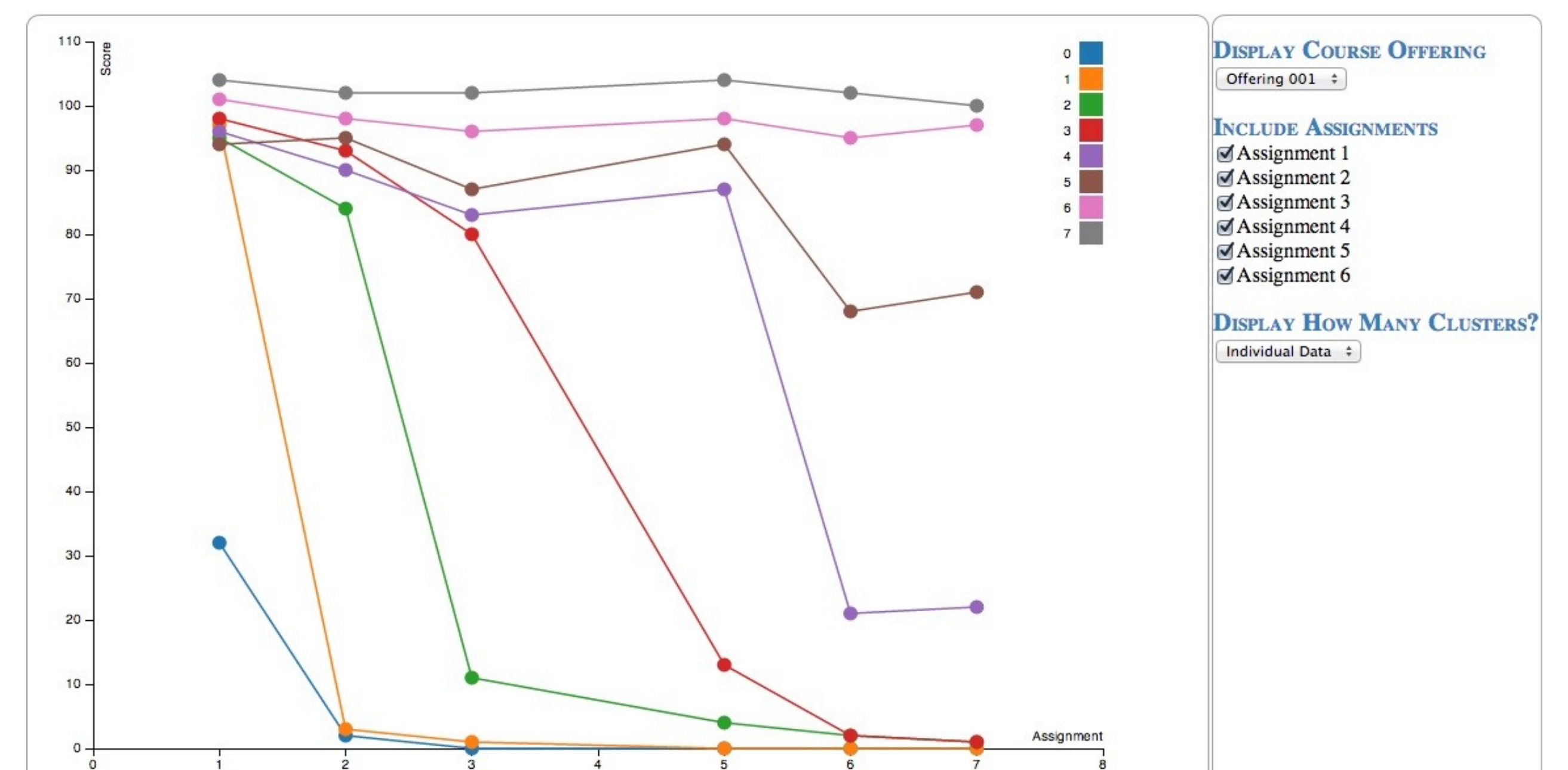
- **Massive fine-grained** course data.
- No existing software/framework specialized in this domain.
- **High demand from instructors** for visualization tools.

## Exam Track



- Slope indicates difficulty levels
- Clustering may indicate common misunderstandings

## Grades



- Instant view comparing how students are trending
- Options to zoom in by **filtering** by assignment

## Approach

### Correlation between statistics

- K-means algorithm to find **best predictor of final grades**

### Comparison across offerings

- Color coded lines/bars for easy comparisons.
- Clustering of individual statistics into bands to **remove outlying data**

### Rich set of visualization options

- Compare **across course offerings**
- Determine **drop-rate** in course participation

## Future Work

### Incorporation of Demographics

- **Filter performance by** background, age group, gender, country of origin etc (privacy preserving)

### Survey Instructors for Feedback

- Determine what elements are commonly included in Coursera offerings
- Discover what **insights and trends** they would like to be able to visualize

### Integration with Coursera Website

- Make **visualizations available by default** to instructors via Coursera or CSV upload