Segmentation of Statistics Canada's Proximity Measures

Weekly Meeting

Week 6

Research Questions

1. What are the optimal cut-off values and cluster boundaries determined by the chosen clustering algorithm in the PMD continuous metric?

2. What distinctive characteristics define each cluster of dissemination blocks, and how do these features contribute to both heterogeneity between clusters and homogeneity within each cluster?

   (Characteristics include: proximity measures, CSD type, DB population, IoR, and province breakdown.)

Methods

- Applied algorithms, analysed clusters, drafted report
  - Manual Cutoffs, OPTICS, Jenks Natural Breaks, Mclust, Kmeans PAM
  - Finished final report methods and results drafting
  - Exploring a new way for cluster validation/no. of clusters

- Sending draft report for feedback early, practice presentation earlier

Equitable Task Distribution

- **Our Approach**
  - Discuss and review the tasks required at our daily group meetings. Every team member is self-motivated to contribute equally.
- **Task Selection**
  - Team members mostly self-selects tasks based on what is needed to advance project. Post discussion, emphasis on moderator rotation and communication.
- **Review & Reflection**
  - At each subsequent meeting we reflect on the task execution, report our findings, and decide next steps
- **Results**
  - Through this method we have maintained an equitable distribution of tasks and dynamic team environment for everyone to contribute and learn

Project Management

- **Team collaboration**
  - Frequent meetings and open discussions have been key in our team dynamic
- **Client Communication**
  - Every member has the opportunity to directly communicate with the client during our weekly meetings. Each of us will present our work results to the client and engage in problem-solving discussions.
- **Meeting Management**
  - Each team member takes turns in creating and moderating the agenda for our meeting.
- **Shared Responsibility**
  - Contribute to the meeting's agenda, and a designated member takes responsibility for moderating the discussion and taking minutes, ensuring active participation.

Team Dynamics

- **Desired Outcome:**
  - **Open communication, talking pain points**
- **What was learned:**
  - **More clear communication, highlight key findings, ‘internal reports’**
  - **Calm down on nitpicking work; ROI, focus on important**
  - **Have better open discussion with team members**
  - **Consensus on adding meeting ie.) after a long day**
  - **Split moderating and management work**

Team Effort - Week 5

Noman

- Mclust model
- K-medoid model
- 28.5 HRS

Ricky

- ‘Manual’ cutoffs
- 31 HRS

Jonah

- Draft final report
- 30.25 HRS

Avishek

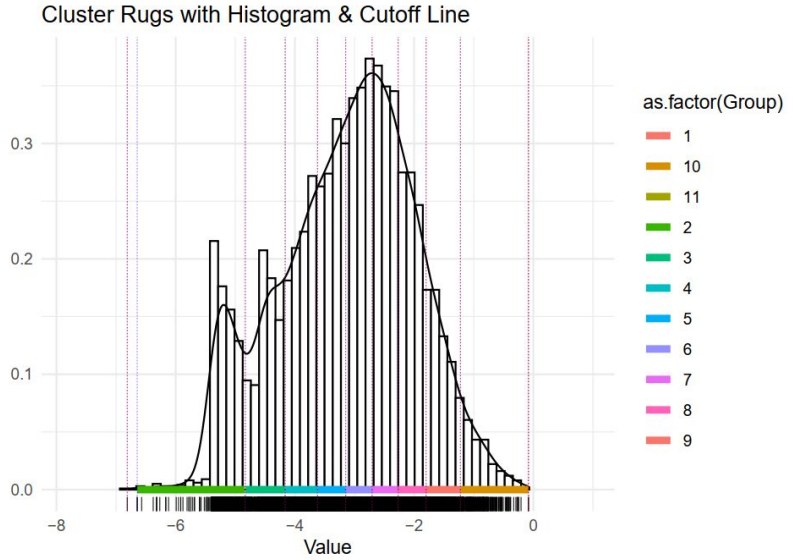
- OPTICS
- Jenks natural breaks
- 26.42 HRS

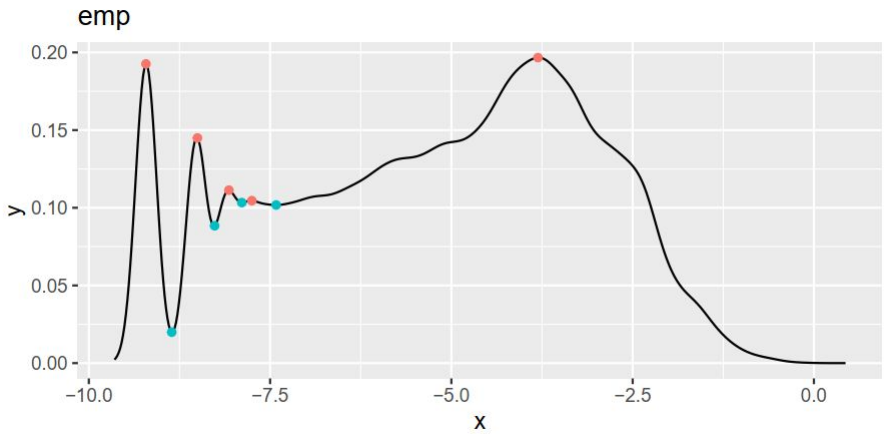
As a Team:

- Team dynamics meeting
- Advanced segmentation methods
- Report start

Progress

- Methods applied
  - Manual Cutoffs
  - Jenks Natural Breaks
  - K-means PAM
    - (Partitioning Around Medoids)
  - Mclust
- Unsuccessful Method
  - OPTICS
  - DENCLUE
- Draft report





Client

- **Client Meeting:** Friday, June 2nd 2023
  - Update on results/findings from the previous week

  - Feedback on things to emphasise/improve for final report & presentation
    - Can compare different algorithms instead of reporting perfect one
    - We should explain theoretical concepts in the final report
    - Compare approaches for robustness check

Upcoming Goals this week

- Cluster Profiling
  - Entire team
- Draft final report
  - Entire team
- Meeting with Jeff
  - Entire team

Upcoming Goals next weeks

- Prepare slides
- Update report based on feedbacks
- Rehearse presentation

Roadblocks/Pivots

- Different metrics gives different results for the same clustering technique.
- Which metric to follow for choosing techniques
- Foresee potential conflict in report writing

Feedback / Questions