

A REPORT  
FOR THE  
SEGMENTATION OF STATISTICS CANADA'S PROXIMITY  
MEASURES

June 2023

Ricky Heinrich  
Noman Mohammad  
Avishek Saha  
Jonah Edmundson

Statistics Canada Liaison - Jérôme Blanchet, Bjenk Ellefsen  
UBCO Project Supervisor - His Excellency Dr. Firas Moosvi

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background</b>	<b>2</b>
<b>3</b>	<b>Data</b>	<b>3</b>
3.1	Primary Dataset . . . . .	3
3.2	Missing Values . . . . .	3
3.3	Other Data . . . . .	3
<b>4</b>	<b>Methods</b>	<b>4</b>
4.1	Data Exploration . . . . .	4
4.2	Clustering Tendency . . . . .	4
4.3	Quintiles . . . . .	4
4.4	Minima Identification . . . . .	4
4.5	Clustering . . . . .	4
4.5.1	Number of Clusters . . . . .	4
4.5.2	Comparison of Algorithms . . . . .	4
4.5.3	Cluster Profiles . . . . .	4
<b>5</b>	<b>Results</b>	<b>5</b>
5.1	Data Exploration . . . . .	5
5.1.1	Summary Statistics . . . . .	5
5.1.2	Distributions . . . . .	5
5.2	Clustering Tendency . . . . .	5
5.3	Quintiles . . . . .	5
5.4	Minima Identification . . . . .	5
5.5	Clustering . . . . .	5
5.5.1	Cluster Profiles . . . . .	5
<b>6</b>	<b>Discussion</b>	<b>6</b>
6.1	Comparison of Approaches . . . . .	6
6.2	Interpretation of Cluster Profiles . . . . .	6
<b>7</b>	<b>Limitations</b>	<b>7</b>
<b>8</b>	<b>Conclusion</b>	<b>8</b>
8.1	Successful Methods . . . . .	8
8.2	Unsuccessful Methods . . . . .	8
8.2.1	Univariate Clustering . . . . .	8
8.2.2	Multivariate Clustering . . . . .	8
8.3	Extra Plots and Tables . . . . .	8
<b>9</b>	<b>References</b>	<b>9</b>
<b>A</b>	<b>Appendix</b>	<b>10</b>

## List of Figures

## List of Tables

# 1 Introduction

The

## 2 Background

## 3 Data

text

### 3.1 Primary Dataset

### 3.2 Missing Values

### 3.3 Other Data

## 4 Methods

text

### 4.1 Data Exploration

### 4.2 Clustering Tendency

### 4.3 Quintiles

### 4.4 Minima Identification

### 4.5 Clustering

#### 4.5.1 Number of Clusters

#### 4.5.2 Comparison of Algorithms

#### 4.5.3 Cluster Profiles



## 5 Results

text

### 5.1 Data Exploration

#### 5.1.1 Summary Statistics

#### 5.1.2 Distributions

### 5.2 Clustering Tendency

### 5.3 Quintiles

### 5.4 Minima Identification

### 5.5 Clustering

#### 5.5.1 Cluster Profiles

## **6 Discussion**

text

### **6.1 Comparison of Approaches**

### **6.2 Interpretation of Cluster Profiles**

## 7 Limitations

## **8 Conclusion**

### **8.1 Successful Methods**

### **8.2 Unsuccessful Methods**

#### **8.2.1 Univariate Clustering**

#### **8.2.2 Multivariate Clustering**

### **8.3 Extra Plots and Tables**

## 9 References

## A Appendix

text