# Cluster Analysis: Identifying Parkinson's Disease Subtypes

Jesse Mu

Wednesday, June 10

# 1 Preprocessing

## 1.1 Dataset Description

## 1.2 Selected Features

Combination of non-motor scale (NMS) symptoms and standard motor symptoms.

| Variable Name | Type | Format | Description |
|---|---|---|---|
| nms_d1 | byte | %8.0g | cardiovascular |
| nms_d2 | byte | %8.0g | sleep/fatigue |
| nms_d3 | byte | %8.0g | mood/cognition |
| nms_d4 | byte | %8.0g | percep/hallucinations |
| nms_d5 | byte | %8.0g | attention/memory |
| nms_d6 | byte | %8.0g | gastrointestinal |
| nms_d7 | byte | %8.0g | urinary |
| nms_d8 | byte | %8.0g | sexual function |
| nms_d9 | byte | %8.0g | miscellaneous |
| tremor | float | %9.0g | tremor |
| bradykin | float | %9.0g | bradykinesia[1] |
| rigidity | float | %9.0g | rigidity |
| axial | float | %9.0g | axial |
| pigd | float | %9.0g | postural instability and gait difficulty |

Table 1: Selected Features and Details

# 2 $k$-means

## 2.1 Identifying optimal number of clusters

No optimal elbow in scree test!

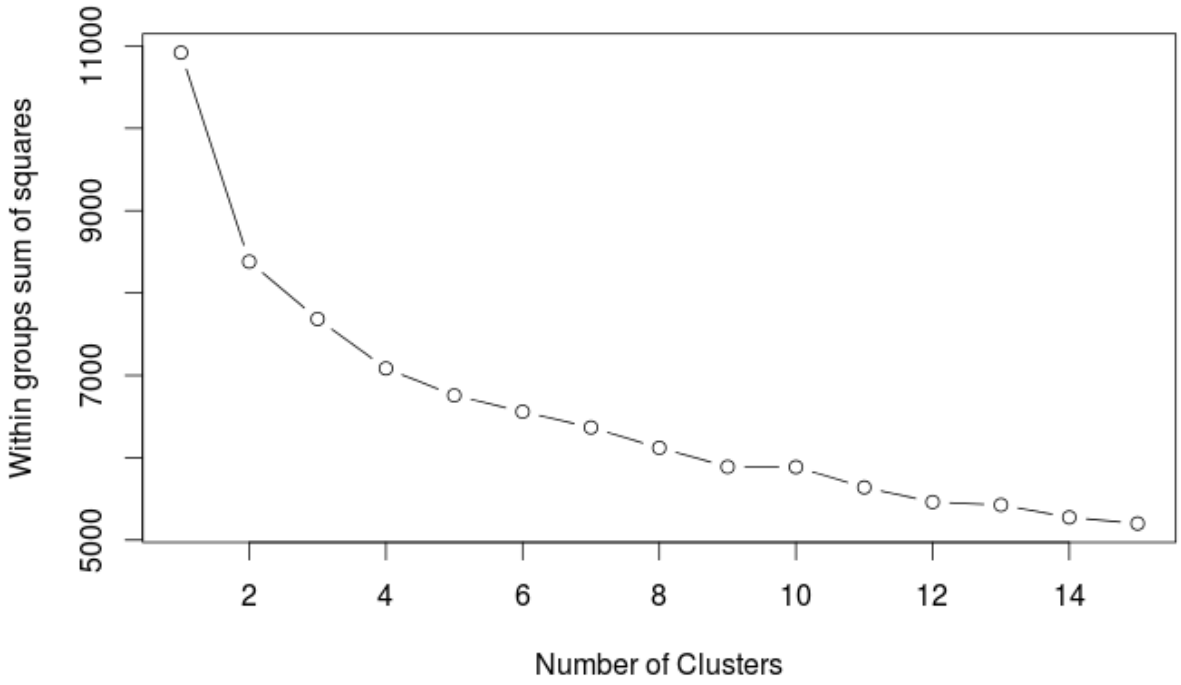


Figure 1: Scree test: WSS error by cluster size