

FML Final Exam

Shiva Chaitanya Goud Gadila

2023-05-03

#IMPORTING OF DATASET

```
PUDL<- read.csv("~/Downloads/fuel_receipts_costs_eia923.csv")
```

#importing the dataset PUDL(Fuel receipts)

#IMPORTING THE LIBRARY

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages — tidyverse  
2.0.0 —
```

```
## ✓ dplyr      1.1.0      ✓ readr      2.1.4
```

```
## ✓ forcats   1.0.0      ✓ stringr   1.5.0
```

```
## ✓ ggplot2    3.4.1      ✓ tibble    3.1.8
```

```
## ✓ lubridate  1.9.2      ✓ tidyr     1.3.0
```

```
## ✓ purrr      1.0.1
```

```
## — Conflicts —
```

```
tidyverse_conflicts() —
```

```
## ✗ dplyr::filter() masks stats::filter()
```

```
## ✗ dplyr::lag()      masks stats::lag()
```

```
## ⓘ Use the ]8;;http://conflicted.r-lib.org/conflicted-package]8;; to force  
all conflicts to become errors
```

```
library(dplyr)
```

```
library(tidyr)
```

```
library(ggplot2)
```

```
library(ggthemes)
```

```
library(caret)
```

```
## Loading required package: lattice
```

```
##
```

```
## Attaching package: 'caret'
```

```
##
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
## lift
```

#Recalling the installed packages.

#CHECKING OF THE DATA

```
str(PUDL)
```

```
## 'data.frame':    608564 obs. of  30 variables:
## $ rowid                : int  1 2 3 4 5 6 7 8 9 10 ...
## $ plant_id_eia         : int  3 3 3 7 7 7 7 8 8 8 ...
## $ plant_id_eia_label   : chr  "Barry" "Barry" "Barry"
"Gadsden" ...
## $ report_date          : chr  "2008-01-01" "2008-01-
01" "2008-01-01" "2008-01-01" ...
## $ contract_type_code   : chr  "C" "C" "C" "C" ...
## $ contract_type_code_label : chr  "C" "C" "C" "C" ...
## $ contract_expiration_date : chr  "2008-04-01" "2008-04-
01" "" "2015-12-01" ...
## $ energy_source_code   : chr  "BIT" "BIT" "NG" "BIT"
...
## $ energy_source_code_label : chr  "BIT" "BIT" "NG" "BIT"
...
## $ fuel_type_code_pudl   : chr  "coal" "coal" "gas"
"coal" ...
## $ fuel_group_code       : chr  "coal" "coal"
"natural_gas" "coal" ...
## $ mine_id_pudl         : int  0 0 NA 1 2 3 NA 4 4 1
...
## $ mine_id_pudl_label   : int  0 0 NA 1 2 3 NA 4 4 1
...
## $ supplier_name        : chr  "interocean coal"
"interocean coal" "bay gas pipeline" "alabama coal" ...
## $ fuel_received_units   : num  259412 52241 2783619
25397 764 ...
## $ fuel_mmbtu_per_unit   : num  23.1 22.8 1.04 24.61
24.45 ...
## $ sulfur_content_pct    : num  0.49 0.48 0 1.69 0.84
1.54 0 2.16 1.24 1.9 ...
## $ ash_content_pct       : num  5.4 5.7 0 14.7 15.5 14.6
0 15.4 11.9 15.4 ...
## $ mercury_content_ppm   : num  NA NA NA NA NA NA NA NA
NA NA ...
## $ fuel_cost_per_mmbtu   : num  2.13 2.12 8.63 2.78 3.38
...
## $ primary_transportation_mode_code : chr  "RV" "RV" "PL" "TR" ...
## $ primary_transportation_mode_code_label : chr  "RV" "RV" "PL" "TR" ...
## $ secondary_transportation_mode_code : chr  "" "" "" "" ...
## $ secondary_transportation_mode_code_label : chr  "" "" "" "" ...
## $ natural_gas_transport_code : chr  "firm" "firm" "firm"
"firm" ...
## $ natural_gas_delivery_contract_type_code : chr  "" "" "" "" ...
## $ moisture_content_pct   : num  NA NA NA NA NA NA NA NA
NA NA ...
## $ chlorine_content_ppm   : num  NA NA NA NA NA NA NA NA
NA NA ...
```

```
## $ data_maturity          : chr  "final" "final" "final"
"final" ...
## $ data_maturity_label    : chr  "final" "final" "final"
"final" ...
```

#EXPLORING THE GIVEN DATA

```
glimpse(PUDL)

## Rows: 608,564
## Columns: 30
## $ rowid                <int> 1, 2, 3, 4, 5, 6, 7, 8,
9, 10...
## $ plant_id_eia         <int> 3, 3, 3, 7, 7, 7, 7, 8,
8, 8,...
## $ plant_id_eia_label   <chr> "Barry", "Barry",
"Barry", "G...
## $ report_date          <chr> "2008-01-01", "2008-01-
01", "...
## $ contract_type_code   <chr> "C", "C", "C", "C", "S",
"S",...
## $ contract_type_code_label <chr> "C", "C", "C", "C", "S",
"S",...
## $ contract_expiration_date <chr> "2008-04-01", "2008-04-
01", "...
## $ energy_source_code    <chr> "BIT", "BIT", "NG",
"BIT", "B...
## $ energy_source_code_label <chr> "BIT", "BIT", "NG",
"BIT", "B...
## $ fuel_type_code_pudl   <chr> "coal", "coal", "gas",
"coal"...
## $ fuel_group_code       <chr> "coal", "coal",
"natural_gas"...
## $ mine_id_pudl         <int> 0, 0, NA, 1, 2, 3, NA, 4,
4, ...
## $ mine_id_pudl_label    <int> 0, 0, NA, 1, 2, 3, NA, 4,
4, ...
## $ supplier_name        <chr> "interoceane coal",
"interoceane...
## $ fuel_received_units   <dbl> 259412, 52241, 2783619,
25397...
## $ fuel_mmbtu_per_unit   <dbl> 23.100, 22.800, 1.039,
24.610...
## $ sulfur_content_pct    <dbl> 0.49, 0.48, 0.00, 1.69,
0.84,...
## $ ash_content_pct       <dbl> 5.4, 5.7, 0.0, 14.7,
15.5, 14...
## $ mercury_content_ppm   <dbl> NA, NA, NA, NA, NA, NA,
NA, N...
## $ fuel_cost_per_mmbtu   <dbl> 2.135, 2.115, 8.631,
```

```

2.776, 3...
## $ primary_transportation_mode_code      <chr> "RV", "RV", "PL", "TR",
"TR",...
## $ primary_transportation_mode_code_label <chr> "RV", "RV", "PL", "TR",
"TR",...
## $ secondary_transportation_mode_code     <chr> "", "", "", "", "", "",
"", "..."
## $ secondary_transportation_mode_code_label <chr> "", "", "", "", "", "",
"", "..."
## $ natural_gas_transport_code             <chr> "firm", "firm", "firm",
"firm..."
## $ natural_gas_delivery_contract_type_code <chr> "", "", "", "", "", "",
"", "..."
## $ moisture_content_pct                  <dbl> NA, NA, NA, NA, NA, NA,
NA, NA...
## $ chlorine_content_ppm                  <dbl> NA, NA, NA, NA, NA, NA,
NA, NA...
## $ data_maturity                         <chr> "final", "final",
"final", "f..."
## $ data_maturity_label                   <chr> "final", "final",
"final", "f..."

```

#Investigating The given Information.

#CLEANING THE GIVEN DATA While the data itself is clean, the dataset contains several variables that have significant missing values. Follows these steps:

#1. Identify all variables that have significant missing values.

col names with missing values


```
colnames(PUDL)[colSums(is.na(PUDL)) > 0]
```

```
## [1] "mine_id_pudl"      "mine_id_pudl_label" "mercury_content_ppm"
## [4] "fuel_cost_per_mmbtu" "moisture_content_pct" "chlorine_content_ppm"
```

all missing values

```
all <- PUDL %>%
  summarise_all(funs(sum(is.na(.)))) %>%
  gather(key = "variable", value = "missing_values") %>%
  filter(missing_values > 0) %>%
  arrange(desc(missing_values))
```

```
## Warning: `funs()` was deprecated in dplyr 0.8.0.
```

```
##  Please use a list of either functions or lambdas:
```

```
##
```

```
## # Simple named list: list(mean = mean, median = median)
```

```
##
```

```
## # Auto named with `tibble::lst()`: tibble::lst(mean, median)
```

```
##
```

```
## # Using lambdas list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
```

```
# Remove variables with significant missing values
PUDL <- PUDL %>%
  select(-all$variable)
```

CHECKING THE DATA

```
str(PUDL)

## 'data.frame':    608564 obs. of  24 variables:
## $ rowid                : int  1 2 3 4 5 6 7 8 9 10 ...
## $ plant_id_eia         : int  3 3 3 7 7 7 7 8 8 8 ...
## $ plant_id_eia_label   : chr  "Barry" "Barry" "Barry"
## "Gadsden" ...
## $ report_date          : chr  "2008-01-01" "2008-01-
01" "2008-01-01" "2008-01-01" ...
## $ contract_type_code   : chr  "C" "C" "C" "C" ...
## $ contract_type_code_label : chr  "C" "C" "C" "C" ...
## $ contract_expiration_date : chr  "2008-04-01" "2008-04-
01" "" "2015-12-01" ...
## $ energy_source_code   : chr  "BIT" "BIT" "NG" "BIT"
...
## $ energy_source_code_label : chr  "BIT" "BIT" "NG" "BIT"
...
## $ fuel_type_code_pudl   : chr  "coal" "coal" "gas"
"coal" ...
## $ fuel_group_code      : chr  "coal" "coal"
"natural_gas" "coal" ...
## $ supplier_name        : chr  "interocean coal"
"interocean coal" "bay gas pipeline" "alabama coal" ...
## $ fuel_received_units   : num  259412 52241 2783619
25397 764 ...
## $ fuel_mmbtu_per_unit   : num  23.1 22.8 1.04 24.61
24.45 ...
## $ sulfur_content_pct    : num  0.49 0.48 0 1.69 0.84
1.54 0 2.16 1.24 1.9 ...
## $ ash_content_pct       : num  5.4 5.7 0 14.7 15.5 14.6
0 15.4 11.9 15.4 ...
## $ primary_transportation_mode_code : chr  "RV" "RV" "PL" "TR" ...
## $ primary_transportation_mode_code_label : chr  "RV" "RV" "PL" "TR" ...
## $ secondary_transportation_mode_code : chr  "" "" "" "" ...
## $ secondary_transportation_mode_code_label : chr  "" "" "" "" ...
## $ natural_gas_transport_code : chr  "firm" "firm" "firm"
"firm" ...
## $ natural_gas_delivery_contract_type_code : chr  "" "" "" "" ...
## $ data_maturity         : chr  "final" "final" "final"
"final" ...
## $ data_maturity_label   : chr  "final" "final" "final"
"final" ...
```

2. Ensure that the variables have the right attributes. For example, numerical or categorical.

```
# attributes
sapply(PUDL, class)

##              rowid
##              "integer"
##              plant_id_eia
##              "integer"
##              plant_id_eia_label
##              "character"
##              report_date
##              "character"
##              contract_type_code
##              "character"
##              contract_type_code_label
##              "character"
##              contract_expiration_date
##              "character"
##              energy_source_code
##              "character"
##              energy_source_code_label
##              "character"
##              fuel_type_code_pudl
##              "character"
##              fuel_group_code
##              "character"
##              supplier_name
##              "character"
##              fuel_received_units
##              "numeric"
##              fuel_mmbtu_per_unit
##              "numeric"
##              sulfur_content_pct
##              "numeric"
##              ash_content_pct
##              "numeric"
##              primary_transportation_mode_code
##              "character"
##              primary_transportation_mode_code_label
##              "character"
##              secondary_transportation_mode_code
##              "character"
##              secondary_transportation_mode_code_label
##              "character"
##              natural_gas_transport_code
##              "character"
##              natural_gas_delivery_contract_type_code
##              "character"
##              data_maturity
```

```
##                                "character"
##                                data_maturity_label
##                                "character"
```

#It determines the data types of each variable in the "PUDL" dataset and returns the information as a vector using the class() function and sapply().

3. To ensure that both the data, and the analysis are unique to each student, randomly sample about 2% of your data using a random 4-digit number as the seed to sample the data. Use 75% of the sampled data as the training set, and the rest as the test set (if needed). This should yield a training set of about 9000 and a test of about 3000.

Set a random seed for reproducibility

```
set.seed(1234)
```

Randomly sample about 2% of the rows from the dataset

```
sampled <- PUDL %>%
  sample_frac(0.02)
```

Split the sampled data into training and test sets

```
train<- sampled %>%
  sample_frac(0.75)
```

```
test<- sampled %>%
  anti_join(train)
```

```
## Joining with `by = join_by(rowid, plant_id_eia, plant_id_eia_label,
## report_date, contract_type_code, contract_type_code_label,
## contract_expiration_date, energy_source_code, energy_source_code_label,
## fuel_type_code_pudl, fuel_group_code, supplier_name, fuel_received_units,
## fuel_mmbtu_per_unit, sulfur_content_pct, ash_content_pct,
## primary_transportation_mode_code, primary_transportation_mode_code_label,
## secondary_transportation_mode_code,
## secondary_transportation_mode_code_label,
## natural_gas_transport_code, natural_gas_delivery_contract_type_code,
## data_maturity, data_maturity_label)`
```

Print the number of rows in the training and test sets

```
nrow(train)
```

```
## [1] 9128
```

```
nrow(test)
```

```
## [1] 3043
```

#This code randomly samples 2% of the "PUDL" dataset, splits it into training and test sets, and prints the number of rows in each set. It is a common data preparation step before building and evaluating predictive models.

#CHECKING THE DATA

```
str(train)
```

```
## 'data.frame':  9128 obs. of  24 variables:
## $ rowid                : int  87571 142756 9625 146942
26617 579028 539024 412250 382869 133924 ...
## $ plant_id_eia         : int  666 2964 55380 1393 2866
7916 57664 50481 2963 4041 ...
## $ plant_id_eia_label   : chr  "J D Kennedy"
"Southwestern" "Union Power Station" "R S Nelson" ...
## $ report_date          : chr  "2009-06-01" "2010-05-
01" "2008-02-01" "2010-06-01" ...
## $ contract_type_code   : chr  "S" "S" "S" "S" ...
## $ contract_type_code_label : chr  "S" "S" "S" "S" ...
## $ contract_expiration_date : chr  "" "" "" "" ...
## $ energy_source_code    : chr  "NG" "NG" "NG" "NG" ...
## $ energy_source_code_label : chr  "NG" "NG" "NG" "NG" ...
## $ fuel_type_code_pudl   : chr  "gas" "gas" "gas" "gas"
...
## $ fuel_group_code       : chr  "natural_gas"
"natural_gas" "natural_gas" "natural_gas" ...
## $ supplier_name        : chr  "florida gas"
"chesapeake" "andarko" "florida gas" ...
## $ fuel_received_units   : num  249079 607 409008 467564
30780 ...
## $ fuel_mmbtu_per_unit   : num  1.06 1.04 1.05 1.03 24.8
...
## $ sulfur_content_pct    : num  0 0 0 0 0.79 0 0 0.95 0
0 ...
## $ ash_content_pct       : num  0 0 0 0 12 0 0 8.7 0 0
...
## $ primary_transportation_mode_code : chr  "" "" "" "" ...
## $ primary_transportation_mode_code_label : chr  "" "" "" "" ...
## $ secondary_transportation_mode_code : chr  "" "" "" "" ...
## $ secondary_transportation_mode_code_label : chr  "" "" "" "" ...
## $ natural_gas_transport_code : chr  "interruptible"
"interruptible" "interruptible" "interruptible" ...
## $ natural_gas_delivery_contract_type_code : chr  "" "" "" "" ...
## $ data_maturity         : chr  "final" "final" "final"
"final" ...
## $ data_maturity_label   : chr  "final" "final" "final"
"final" ...
```

```
str(test)
```

```
## 'data.frame':  3043 obs. of  24 variables:
## $ rowid                : int  126055 382554 345167
199608 279106 237360 330424 131974 166742 413590 ...
## $ plant_id_eia         : int  50978 1733 3399 55192 96
6061 8102 535 8 2723 ...
## $ plant_id_eia_label   : chr  "Carr Street" "Monroe"
```



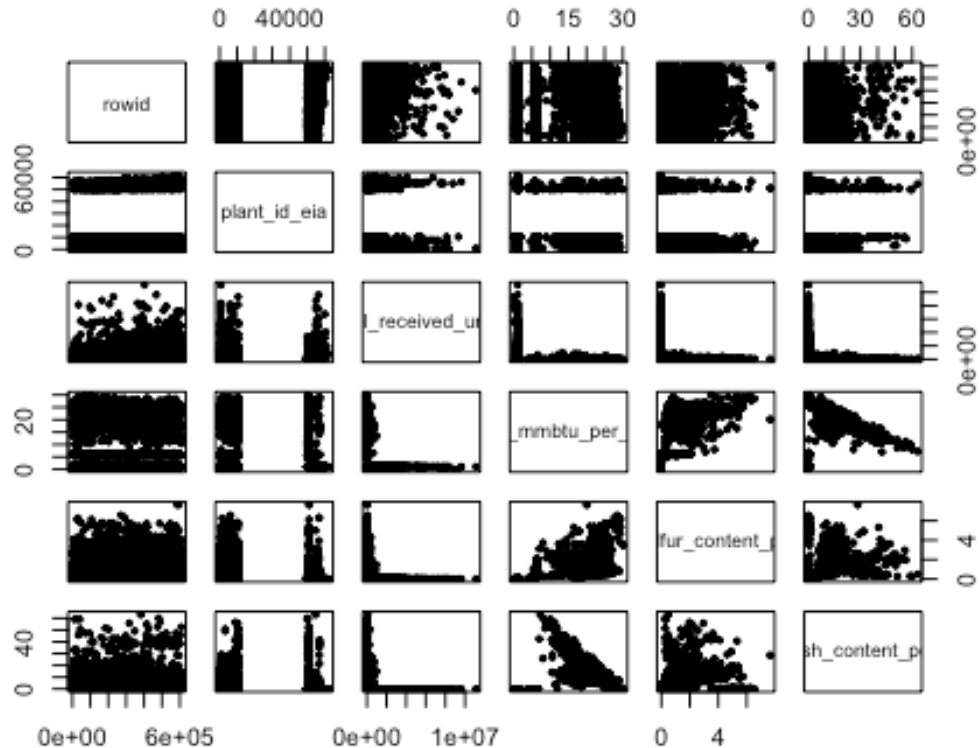
```

"Cumberland" "Osceola" ...
## $ report_date : chr "2010-01-01" "2015-11-01" "2014-10-01" "2011-04-01" ...
## $ contract_type_code : chr "S" "C" "S" "S" ...
## $ contract_type_code_label : chr "S" "C" "S" "S" ...
## $ contract_expiration_date : chr "" "2015-11-01" "" ""
...
## $ energy_source_code : chr "NG" "BIT" "DFO" "NG"
...
## $ energy_source_code_label : chr "NG" "BIT" "DFO" "NG"
...
## $ fuel_type_code_pudl : chr "gas" "coal" "oil" "gas"
...
## $ fuel_group_code : chr "natural_gas" "coal"
"petroleum" "natural_gas" ...
## $ supplier_name : chr "sprague energy corp"
"blackhawk mining llc" "jat oil" "seminole" ...
## $ fuel_received_units : num 11537 12883 170 163405
875779 ...
## $ fuel_mmbtu_per_unit : num 1.03 25.1 5.76 1.03 1
...
## $ sulfur_content_pct : num 0 0.76 0 0 0 0.84 3.8 0
0.99 0 ...
## $ ash_content_pct : num 0 8.2 0 0 0 ...
## $ primary_transportation_mode_code : chr "PL" "RR" "TR" "PL" ...
## $ primary_transportation_mode_code_label : chr "PL" "RR" "TR" "PL" ...
## $ secondary_transportation_mode_code : chr "" "" "" "" ...
## $ secondary_transportation_mode_code_label : chr "" "" "" "" ...
## $ natural_gas_transport_code : chr "interruptible" "" ""
"firm" ...
## $ natural_gas_delivery_contract_type_code : chr "" "" "" "" ...
## $ data_maturity : chr "final" "final" "final"
"final" ...
## $ data_maturity_label : chr "final" "final" "final"
"final" ...

# Identify the numeric variables in the training set
numVars <- names(train)[sapply(train, is.numeric)]

# Create the scatterplot matrix using the pairs() function
pairs(train[,numVars], pch = 19, cex = 0.5,)

```



#The code selects the numeric variables from the train dataframe and stores them in numVars. It then creates a scatterplot matrix using the pairs() function to visualize the pairwise relationships between the numeric variables in the train dataset.

k-means clustering

```
set.seed(1234)
```

Select numeric columns from the train data

```
numValues <- train %>% select_if(is.numeric)
```

Perform k-means clustering with 3 clusters

```
kmeans <- kmeans(numValues, centers = 3)
```

```
kmeans
```

```
## K-means clustering with 3 clusters of sizes 583, 111, 8434
```

```
##
```

```
## Cluster means:
```

```
##      rowid plant_id_eia fuel_received_units fuel_mmbtu_per_unit
```

```
## 1 353664.9   37772.14      1734967.04         1.0507136
```

```
## 2 382242.5   28576.76      5018917.32         0.9681261
```

```
## 3 301395.7   16623.99       81635.17         9.4080154
```

```
##      sulfur_content_pct ash_content_pct
```

```
## 1      0.002521441      0.02504288
```

[illegible]

[illegible]

```
## [1777] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3
3 1 3
## [1814] 3 3 3 3 3 3 3 3 1 1 3 3 3 1 3 3 3 3 3 2 3 3 1 3 3 3 3 3 3 3 3 3
3 3 3
## [1851] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 1 3 3
3 3 3
## [1888] 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 1 3 3 3 3
1 3 3
## [1925] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 1 3 3 3 3
3 3 3
## [1962] 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 3 3
3 3 3
## [1999] 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [2036] 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3
3 3 3
## [2073] 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3
3 3 3
## [2110] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 2 3 1 3 3 1 1 3 3 3 3 3 2 3 3
3 3 3
## [2147] 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 2 3 3 3 3
1 3 1
## [2184] 3 3 3 3 3 3 1 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3
3 3 3
## [2221] 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [2258] 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 2 1
3 3 3
## [2295] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [2332] 3 3 3 3 3 3 3 1 3 3 3 2 1 1 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 1 3 3 3
3 3 3
## [2369] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1
3 3 3
## [2406] 3 3 3 3 3 1 3 1 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 2
1 3 3
## [2443] 3 3 1 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 1 3 1 3 3 3 3 3 1 3 1
3 3 3
## [2480] 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 2 1 3 3 1 3
3 3 3
## [2517] 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 1 3 3 3 3 3 3
3 3 3
## [2554] 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 1 3 3 1 3 3 3 3 3 3 3
1 3 3
## [2591] 1 1 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3
3 3 3
## [2628] 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [2665] 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
```

[illegible]

[illegible]

[illegible]

[illegible]

[6402] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 2 1 3 3 3 3 3 3 3 2 3 3 3 1 3
3 3 3
[6439] 2 3 3 3 3 3 3 1 3 1 3 1 3 3 1 3 3 3 3 3 3 3 3 1 3 3 3 2 1 3 3 3 3 3
3 3 3
[6476] 2 3 3 3 3 2 1 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1
3 1 3
[6513] 3 3 3 3 3 3 1 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 1 3
3 3 3
[6550] 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3
3 3 3
[6587] 3 3 3 2 1 3 1 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3
3 3 3
[6624] 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[6661] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
1 3 1
[6698] 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[6735] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 2 3 3 3
3 3 3
[6772] 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3
3 3 3
[6809] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3
3 3 3
[6846] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3
3 3 3
[6883] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 3 3 3 1 3 3 3 3 3 3 3 3 3 3
3 3 1
[6920] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 1
3 3 3
[6957] 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 1 3 3 3 3 3 1 3 3
3 1 3
[6994] 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 1 3 2 3 3 3
3 3 3
[7031] 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[7068] 3 3 3 3 3 3 3 3 3 3 3 1 1 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3
3 3 3
[7105] 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3
3 3 3
[7142] 3 3 3 3 2 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[7179] 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[7216] 3 3 3 3 3 3 1 3 3 1 3 3 3 3 3 3 3 2 3 3 3 3 1 3 3 3 1 3 3 3 3 3 3 3
3 3 3
[7253] 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3
3 3 1
[7290] 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3

[illegible]


```

## (between_SS / total_SS = 79.6 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"

#It performs k-means clustering with 3 clusters on the numeric columns of the training data. The resulting clusters will be stored in the kmeans object.

# aggregate the data
aggregate(train[,numVars], by = list (kmeans $ cluster), mean)

##   Group.1   rowid plant_id_eia fuel_received_units fuel_mmbtu_per_unit
## 1      1 353664.9   37772.14         1734967.04         1.0507136
## 2      2 382242.5   28576.76         5018917.32         0.9681261
## 3      3 301395.7   16623.99          81635.17         9.4080154
##   sulfur_content_pct ash_content_pct
## 1      0.002521441    0.02504288
## 2      0.000000000    0.00000000
## 3      0.558086317    3.79857482

#This code aggregates the numeric variables in the training set by the k-means cluster labels and computes the mean for each variable within each cluster.

#visualizing the given data

# Create a ggplot object with the train data and cluster assignments on the y-axis
ggplot(train, aes(y = kmeans$cluster)) +

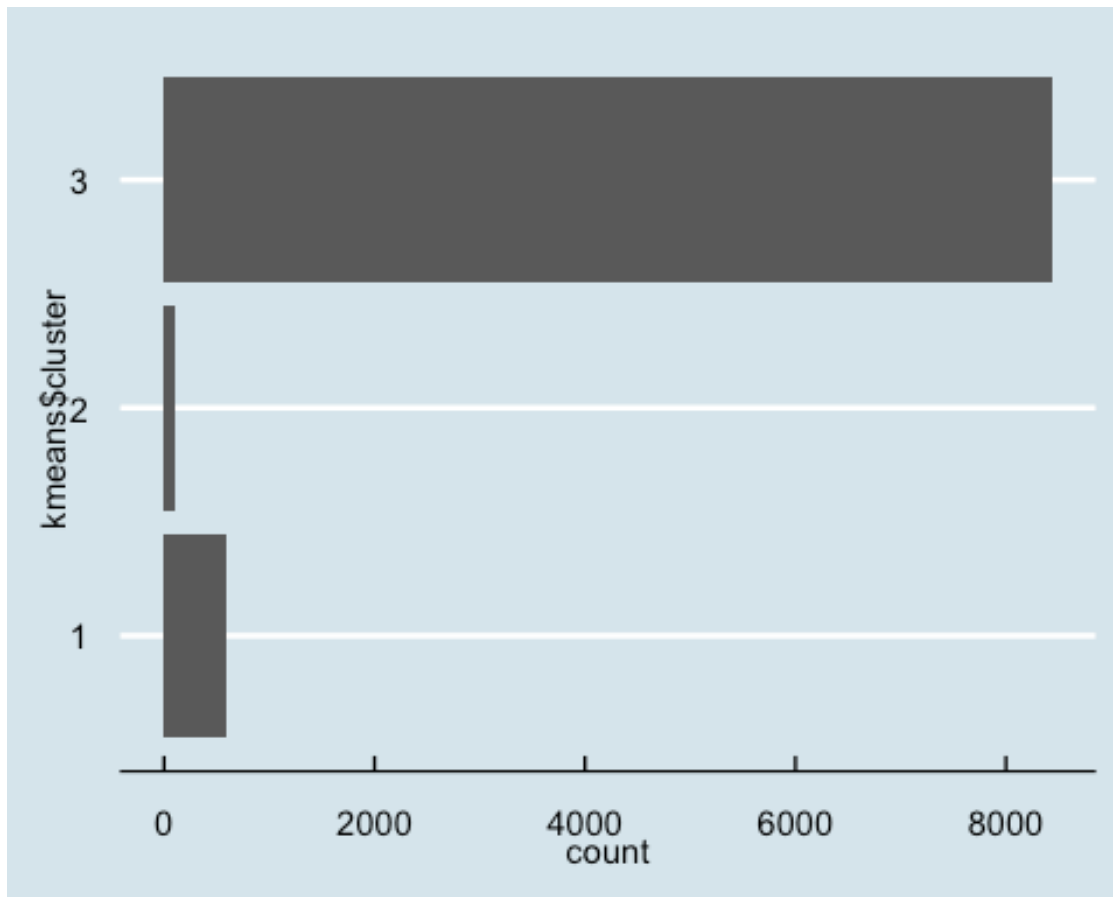
# Add a bar layer to the plot, with fill color based on cluster assignments
  geom_bar(aes(fill = kmeans$cluster), position = "dodge") +

# Apply the economist theme to the plot
  theme_economist() +

# Adjust the title alignment
  theme(plot.title = element_text(hjust = 0.5))

## Warning: The following aesthetics were dropped during statistical transformation: fill
## ⓘ This can happen when ggplot fails to infer the correct grouping structure in the data.
## ⓘ Did you forget to specify a `group` aesthetic or to convert a numerical variable into a factor?

```



#This creates a bar plot using ggplot, with the cluster assignments on the y-axis and the fill color based on the cluster assignments. The economist theme is applied to the plot, and the title alignment is adjusted.

#KNN

```
# Set the seed for reproducibility
set.seed(1234)

# Identify numeric columns in the train data
numValues <- sapply(train, is.numeric)

# Load the 'class' package for kNN classification
library(class)

#Subset the train and test data to only include numeric columns
train1 <- train[, numValues]
test1 <- test[, numValues]

# Perform kNN classification on the numeric train and test data
# using the cluster assignments from the k-means model as the class labels
knn <- knn(train1, test1, cl = kmeans$cluster, k = 3)
knn
```

```
## [1] 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [38] 3 3 3 3 3 3 3 3 1 3 3 3 3 1 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1
3 3 3
## [75] 1 3 3 3 3 3 3 3 1 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 2
## [112] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3
1 3 3
## [149] 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 1 3 1 3 3 3 3 3 3 3 3
3 1 3
## [186] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2
3 3 3
## [223] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [260] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3
3 1 3
## [297] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [334] 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [371] 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 1 1 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [408] 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 1 3 3 3 3 3 3 3 3
1 1 3
## [445] 3 3 1 3 3 3 3 3 3 1 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2
3 3 3
## [482] 1 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
1 2 3
## [519] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3
3 3 3
## [556] 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
1 3 3
## [593] 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [630] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 1 3 3 1 1 3 3 3 3 3 3 3 3 3
1 3 3
## [667] 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 1 3 3 2 3 1 3 3 3 1 3 3
3 3 3
## [704] 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [741] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3
3 3 3
## [778] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 3 3 1
3 3 3
## [815] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [852] 3 3 3 1 3 3 3 3 3 3 1 1 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 1
3 2 3
## [889] 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 1 3
```

```
## [926] 3 3 3 3 3 3 3 1 1 3 1 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3
3 3 3
## [963] 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3
3 3 3
## [1000] 3 3 3 1 3 3 3 3 3 3 1 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1037] 3 3 3 3 3 3 3 3 3 1 3 3 3 3 1 3 1 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 2
3 3 3
## [1074] 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
1 3 3
## [1111] 3 3 1 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3
3 3 3
## [1148] 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 2 3 3 3 3 1 3 3 3 3
3 3 3
## [1185] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 2 3 3 3 3
3 3 3
## [1222] 3 2 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1259] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 1 3 3
3 3 3
## [1296] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 1 3 3 3 3
3 3 3
## [1333] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 1 3 3 3 3
3 3 3
## [1370] 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3
3 3 3
## [1407] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 2 3 3 3 3 3 1 3 3 3 3
3 3 3
## [1444] 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1481] 3 3 3 3 3 3 3 3 3 3 1 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3
3 3 3
## [1518] 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 1 1 3 3 3 3 3 1
3 3 3
## [1555] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1592] 3 3 3 3 3 3 3 3 2 3 3 3 3 1 3 3 1 3 3 1 3 3 1 3 3 3 3 3 3 3 3 3
3 3 3
## [1629] 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1666] 3 3 3 3 3 1 3 3 3 3 3 3 2 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1703] 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1740] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 1 3 3 3 3 3 3 3
3 3 3
## [1777] 3 3 3 3 3 3 3 1 3 3 1 2 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3
3 3 3
## [1814] 3 3 3 1 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 2
3 3 2
```


[1851] 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 1 3 3 3 3 3 1 3 3 3 3
3 3 3
[1888] 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1
3 3 3
[1925] 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[1962] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 3 1 3 3 3 3 3 3 3 3 3
3 3 3
[1999] 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3
3 1 3
[2036] 3 3 3 3 3 3 3 3 3 3 3 1 3 3 1 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3
3 3 3
[2073] 1 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 1
[2110] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[2147] 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 1 3 3 3 3 3 1 3 3 3 3 3 3 3 3
3 3 3
[2184] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 1 3 3 3 1 3 3 3 1 3 3 3
3 3 3
[2221] 3 3 3 3 3 3 3 3 3 3 2 3 3 2 3 3 3 3 1 3 3 3 3 3 1 3 3 3 3 3 3
3 3 3
[2258] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[2295] 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3
3 1 3
[2332] 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[2369] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3
3 3 3
[2406] 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3
3 3 3
[2443] 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 1
3 3 3
[2480] 3 1 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 1 3 3 3 1 3 3 3 3 3 3 3
3 3 3
[2517] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1
3 1 3
[2554] 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 1 3 3 3 3 3 3 3 3 3
3 3 3
[2591] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[2628] 3 3 3 3 3 3 3 3 3 1 3 3 1 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 1 3 1
3 3 3
[2665] 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
[2702] 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3
3 3 2
[2739] 3 3 1 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3
3 3 3

```
## [2776] 3 2 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 3 3
## [2813] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1 3 3 3 3
1 3 3
## [2850] 3 3 3 3 2 3 3 3 1 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
3 1 1
## [2887] 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 1 3 3 3 3 3
3 3 3
## [2924] 3 3 1 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 1 3 3 3 3
3 1 3
## [2961] 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3
3 3 3
## [2998] 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 1 3 3 3 3 3 3 3 1 3 3 3 3 3 3 1
3 3 3
## [3035] 3 3 3 3 3 3 3 3 3
## Levels: 1 2 3
```

#It perform k-Nearest Neighbors classification on the numeric columns of the train and test data, using the cluster assignments obtained from k-means clustering with 3 clusters on the train data as the class labels. The resulting classification is stored in the knn object.

#SEGMENTATION

#Identify numeric columns in the train data

```
numValues <- sapply(train, is.numeric)
```

Step 2: Perform k-means clustering on the numeric train data, with 3 clusters

```
kmeans <- kmeans(train[, numValues], centers = 3)
```

```
kmeans
```

```
## K-means clustering with 3 clusters of sizes 583, 8434, 111
```

```
##
```

```
## Cluster means:
```

```
##      rowid plant_id_eia fuel_received_units fuel_mmbtu_per_unit
```

```
## 1 353664.9    37772.14        1734967.04         1.0507136
```

```
## 2 301395.7    16623.99         81635.17         9.4080154
```

```
## 3 382242.5    28576.76        5018917.32         0.9681261
```

```
##      sulfur_content_pct ash_content_pct
```

```
## 1      0.002521441      0.02504288
```

```
## 2      0.558086317      3.79857482
```

```
## 3      0.000000000      0.00000000
```

```
##
```

```
## Clustering vector:
```

```
##      [1] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 1 1 2 2 1 2 2 2 2 2
2 1 2
```

```
##      [38] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2
2 2 1
```

```
##      [75] 2 2 2 2 2 2 2 2 2 3 2 2 2 1 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2
```

[illegible]

```

2 2 2
## [1037] 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 1 1 2 2 2 2 2 2 2
2 2 2
## [1074] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1
2 2 2
## [1111] 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2
2 2 2
## [1148] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 1 2 2 2 2 2 2 2 2 2
2 3 2
## [1185] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
1 2 2
## [1222] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2
2 2 2
## [1259] 2 2 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2
2 2 2
## [1296] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
## [1333] 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2
2 2 2
## [1370] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 1 2
## [1407] 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
## [1444] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2
2 2 2
## [1481] 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
## [1518] 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2
2 2 2
## [1555] 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 3 2 2 2 1 2
2 2 2
## [1592] 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2 2 2
2 2 2
## [1629] 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
1 2 2
## [1666] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 1 2 2 1 2 2
2 2 2
## [1703] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2
2 2 2
## [1740] 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
## [1777] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2
2 1 2
## [1814] 2 2 2 2 2 2 2 2 1 1 2 2 2 1 2 2 2 2 3 2 2 1 2 2 2 2 2 2 2
2 2 2
## [1851] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 2
2 2 2
## [1888] 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 1 2
1 2 2
## [1925] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 1 2 2 2

```

2 2 2
[1962] 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2
2 2 2
[1999] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[2036] 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 3 2 2 2
2 2 2
[2073] 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2
2 2 2
[2110] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 3 2 1 2 2 1 1 2 2 2 2 2 3 2 2
2 2 2
[2147] 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 3 2 2 2 2
1 2 1
[2184] 2 2 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2
2 2 2
[2221] 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[2258] 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 3 1
2 2 2
[2295] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[2332] 2 2 2 2 2 2 2 1 2 2 2 3 1 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2 2
2 2 2
[2369] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1
2 2 2
[2406] 2 2 2 2 2 1 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 3
1 2 2
[2443] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 2 1 2 2 2 2 2 1 2 1
2 2 2
[2480] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 3 1 2 2 1 2
2 2 2
[2517] 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2 2 2 2
2 2 2
[2554] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2 1 2 2 2 2 2
1 2 2
[2591] 1 1 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2
2 2 2
[2628] 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[2665] 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[2702] 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2 2 2 2 1 3 1 2 2 2 2 2 2 2 2 2 2
1 2 2
[2739] 2 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2
2 1 2
[2776] 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 3 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2
1 2 2
[2813] 2 2 2 2 2 3 3 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2
2 2 2
[2850] 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2

2 2 2
[3812] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2
2 2 2
[3849] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[3886] 2 2 3 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 1 2 2 2 2 2 2 2
2 2 2
[3923] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 2 2 2 2 2 2 2
2 1 2
[3960] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[3997] 1 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2
2 2 2
[4034] 2 2 2 2 2 3 2 2 2 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4071] 1 2 2 2 2 2 1 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4108] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4145] 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 1 1 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4182] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2
2 2 2
[4219] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4256] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4293] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4330] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4367] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 1 2 2 2 2 2 1 1 2 2 2 2 2 2
2 2 2
[4404] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 1 2 3 2 2 2 2 2 2 2 2 2 2 2
1 2 2
[4441] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 1 3 2 2 2 2 2 2
2 1 2
[4478] 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4515] 1 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4552] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4589] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4626] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4663] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2
2 2 2
[4700] 2 2 2 2 2 1 2 2 2 1 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2

2 1 2
[4737] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2
2 2 2
[4774] 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4811] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4848] 1 2 2 2 2 2 2 1 1 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2
2 2 2
[4885] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2
2 2 2
[4922] 2 2 2 2 1 2 2 1 2 2 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[4959] 2 2 3 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 2 2
2 2 2
[4996] 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5033] 2 2 2 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 3 1 2 2 2 2 2 2 2 2 2
2 2 2
[5070] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 1 1 2 2 2 2 2 2
2 2 2
[5107] 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5144] 2 2 2 2 2 3 1 2 2 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 1 2 2
2 2 2
[5181] 2 2 2 1 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 2
2 2 2
[5218] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 1 2
[5255] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5292] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5329] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 1 2 1 2 2 2 2 2 2
2 2 2
[5366] 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5403] 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5440] 2 2 2 2 2 2 2 2 2 2 2 2 2 3 1 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2
2 2 2
[5477] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2
2 2 2
[5514] 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 1 2 2 2 2 2 2
2 2 2
[5551] 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5588] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2
2 2 2
[5625] 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2

2 2 2
[5662] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2
2 2 2
[5699] 2 2 2 2 2 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2
1 2 2
[5736] 2 2 1 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 3 2 2 3 2 2 2 2 2 2 2 2 2 2
2 2 2
[5773] 2 2 2 1 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2
2 2 2
[5810] 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 3 2 2 2 1 2 2 3 2 1 2 2 2
2 2 2
[5847] 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 1 2 2 2 2
2 2 2
[5884] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2
2 2 2
[5921] 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[5958] 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2
2 2 2
[5995] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2 2
2 2 2
[6032] 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6069] 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6106] 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6143] 2 2 2 2 2 3 2 2 2 2 2 2 3 2 2 2 1 2 2 2 1 1 2 2 1 2 2 2 2 2 2 2 2 2
2 2 2
[6180] 2 2 1 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6217] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 1
[6254] 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 3 2 2 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6291] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2
2 2 2
[6328] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6365] 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6402] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 3 1 2 2 2 2 2 2 2 3 2 2 2 1 2
2 2 2
[6439] 3 2 2 2 2 2 2 1 2 1 2 1 2 2 1 2 2 2 2 2 2 2 1 2 2 2 3 1 2 2 2 2 2
2 2 2
[6476] 3 2 2 2 2 3 1 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1
2 1 2
[6513] 2 2 2 2 2 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 1 2
2 2 2
[6550] 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2

2 2 2
[6587] 2 2 2 3 1 2 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2
2 2 2
[6624] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6661] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
1 2 1
[6698] 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[6735] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 3 2 2 2
2 2 2
[6772] 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2
2 2 2
[6809] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2
2 2 2
[6846] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2
2 2 2
[6883] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2
2 2 1
[6920] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 1
2 2 2
[6957] 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 2 2 2 2 2 1 2 2
2 1 2
[6994] 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 3 2 2 2
2 2 2
[7031] 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[7068] 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2
2 2 2
[7105] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2
2 2 2
[7142] 2 2 2 2 3 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[7179] 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[7216] 2 2 2 2 2 2 1 2 2 1 2 2 2 2 2 2 2 2 3 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 2
2 2 2
[7253] 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2
2 2 1
[7290] 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[7327] 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[7364] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2
2 2 2
[7401] 2 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2
2 2 2
[7438] 2 2 2 1 2 2 2 2 2 1 3 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2
2 2 2
[7475] 2 2 2 2 2 3 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2

[illegible]

```

2 2 1
## [8437] 2 2 2 2 2 2 2 1 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
## [8474] 2 2 2 2 2 2 3 1 2 2 2 2 2 1 2 1 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2
2 2 2
## [8511] 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 3 2 2 2
2 2 2
## [8548] 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
## [8585] 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2
2 2 2
## [8622] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 2 2 2 2 2 2 2 2 2 2
2 2 2
## [8659] 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 2
## [8696] 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2
2 2 1
## [8733] 2 2 2 2 2 2 2 2 1 1 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2
2 1 2
## [8770] 2 2 2 2 2 2 2 2 2 1 2 2 2 2 3 2 2 3 2 2 2 1 2 2 2 2 2 2 2 2 1 2
2 2 2
## [8807] 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2
2 2 2
## [8844] 2 2 2 2 2 2 2 2 3 2 1 2 1 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 2 1
2 2 2
## [8881] 2 2 2 2 2 2 2 2 2 3 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 1 2 2 1 1 2 2
2 2 2
## [8918] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 1 2 2 2
2 2 2
## [8955] 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1
2 2 2
## [8992] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2
2 2 2
## [9029] 2 2 2 1 2 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2
2 2 2
## [9066] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2
2 2 2
## [9103] 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2
##
## Within cluster sum of squares by cluster:
## [1] 2.670057e+14 4.634426e+14 3.078375e+14
## (between_SS / total_SS = 79.6 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"

```

#This code performs k-means clustering on the numeric columns of the training data with 3 clusters. The resulting object is saved as kmeans.

Step 1: Get the length of the k-means cluster assignments

```
kmeans_length <- length(kmeans$cluster)
```

```
kmeans_length
```

```
## [1] 9128
```

Step 2: Get the length of the kNN classifications

```
knn_length <- length(knn)
```

```
knn_length
```

```
## [1] 3043
```

It calculates the length of the k-means cluster assignments and the length of the kNN classifications, stored in the variables kmeans_length and knn_length, respectively.