

Report on my own edx project

FUNG CHE HEI

8/19/2020

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
## Loading required package: tidyverse
```

```
## -- Attaching packages ----- tidyverse_
```

```
## v ggplot2 3.3.2    v purrr  0.3.4
## v tibble  3.0.3    v dplyr  1.0.1
## v tidyr   1.1.1    v stringr 1.4.0
## v readr   1.3.1    v forcats 0.5.0
```

```
## -- Conflicts ----- tidyverse_conf
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
## Loading required package: caret
```

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

```
##
```

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

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

```
##
```

```
## lift
```

```
## Loading required package: data.table
```

```
##
```

```
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
```

```
##
```

```
## between, first, last
```

```

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

## Loading required package: dslabs

## Loading required package: lubridate

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:data.table':
##
##      hour, isoweek, mday, minute, month, quarter, second, wday, week,
##      yday, year

## The following objects are masked from 'package:base':
##
##      date, intersect, setdiff, union

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   seismic = col_character(),
##   seismoacoustic = col_character(),
##   shift = col_character(),
##   ghazard = col_character()
## )

## See spec(...) for full column specifications.

##           id           seismic      seismoacoustic      shift
## Min.      : 1.0   Length:2584   Length:2584   Length:2584
## 1st Qu.: 646.8   Class :character   Class :character   Class :character
## Median :1292.5   Mode  :character   Mode  :character   Mode  :character
## Mean      :1292.5
## 3rd Qu.:1938.2
## Max.      :2584.0
##      genergy      gpuls      gdenergy      gdpuls
## Min.      : 100   Min.      : 2.0   Min.      : -96.00   Min.      : -96.000
## 1st Qu.: 11660   1st Qu.: 190.0   1st Qu.: -37.00   1st Qu.: -36.000
## Median : 25485   Median : 379.0   Median : -6.00   Median : -6.000
## Mean      : 90242   Mean      : 538.6   Mean      : 12.38   Mean      : 4.509
## 3rd Qu.: 52832   3rd Qu.: 669.0   3rd Qu.: 38.00   3rd Qu.: 30.250
## Max.      :2595650   Max.      :4518.0   Max.      :1245.00   Max.      :838.000
##      ghazard      nbumps      nbumps2      nbumps3
## Length:2584   Min.      :0.0000   Min.      :0.0000   Min.      :0.0000
## Class :character   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000
## Mode  :character   Median :0.0000   Median :0.0000   Median :0.0000
##                      Mean      :0.8595   Mean      :0.3936   Mean      :0.3928
##                      3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000
##                      Max.      :9.0000   Max.      :8.0000   Max.      :7.0000

```

```

##      nbumps4      nbumps5      nbumps6      nbumps7      nbumps89
## Min.   :0.00000   Min.   :0.000000   Min.   :0   Min.   :0   Min.   :0
## 1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:0   1st Qu.:0   1st Qu.:0
## Median :0.00000   Median :0.000000   Median :0   Median :0   Median :0
## Mean   :0.06772   Mean   :0.004644   Mean   :0   Mean   :0   Mean   :0
## 3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.:0   3rd Qu.:0   3rd Qu.:0
## Max.   :3.00000   Max.   :1.000000   Max.   :0   Max.   :0   Max.   :0
##      energy      maxenergy      class
## Min.   :      0   Min.   :      0   Min.   :0.00000
## 1st Qu.:      0   1st Qu.:      0   1st Qu.:0.00000
## Median :      0   Median :      0   Median :0.00000
## Mean   :  4975   Mean   :  4279   Mean   :0.06579
## 3rd Qu.: 2600   3rd Qu.: 2000   3rd Qu.:0.00000
## Max.   :402000   Max.   :400000   Max.   :1.00000

##      id seismic seismoacoustic shift genenergy gpuls gdenenergy gdpuls ghazard nbumps
## 1  1      a              a      N   15180    48      -72    -72      a      0
## 2  2      a              a      N   14720    33      -70    -79      a      1
## 3  3      a              a      N    8050    30      -81    -78      a      0
## 4  4      a              a      N   28820   171      -23     40      a      1
## 5  5      a              a      N   12640    57      -63    -52      a      0
## 6  6      a              a      W   63760   195      -73    -65      a      0
##      nbumps2 nbumps3 nbumps4 nbumps5 nbumps6 nbumps7 nbumps89 energy maxenergy
## 1      0      0      0      0      0      0      0      0      0      0
## 2      0      1      0      0      0      0      0      2000    2000
## 3      0      0      0      0      0      0      0      0      0      0
## 4      0      1      0      0      0      0      0      3000    3000
## 5      0      0      0      0      0      0      0      0      0      0
## 6      0      0      0      0      0      0      0      0      0      0
##      class
## 1      0
## 2      0
## 3      0
## 4      0
## 5      0
## 6      0

```

Introduction

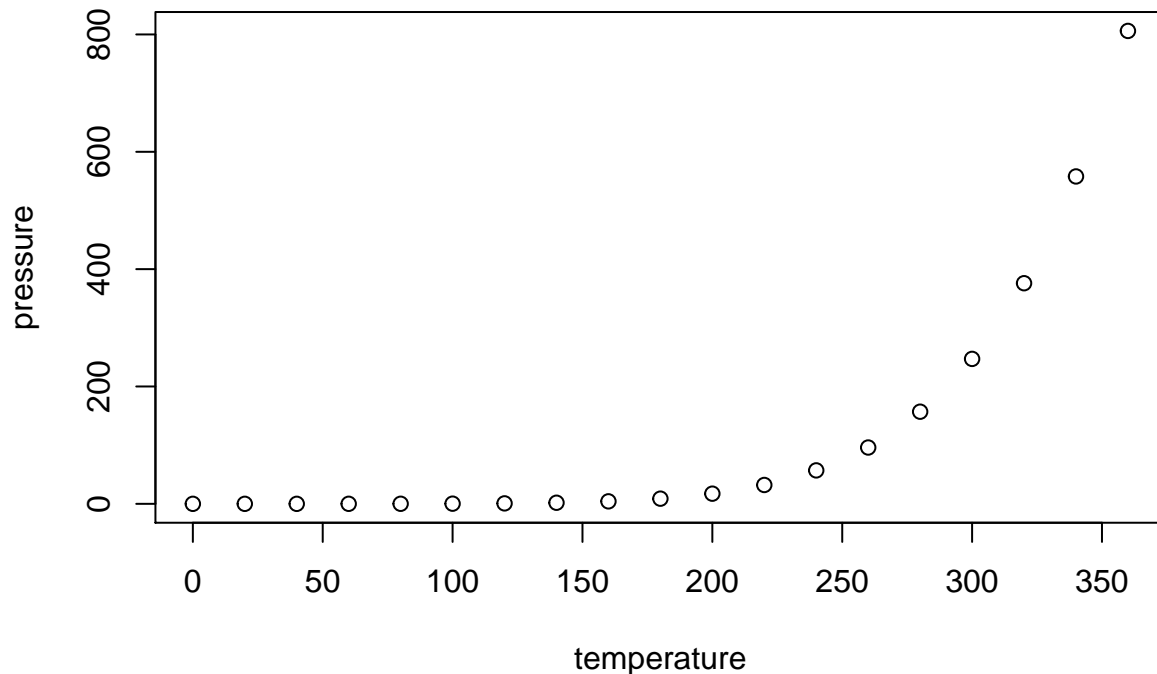
Attribute Information

1. seismic: result of shift seismic hazard assessment in the mine working obtained by the seismic method (a - lack of hazard, b - low hazard, c - high hazard, d - danger state);
2. seismoacoustic: result of shift seismic hazard assessment in the mine working obtained by the seismoacoustic method;
3. shift: information about type of a shift (W - coal-getting, N -preparation shift);
4. genenergy: seismic energy recorded within previous shift by the most active geophone (GMax) out of geophones monitoring the longwall;
5. gpuls: a number of pulses recorded within previous shift by GMax;
6. gdenenergy: a deviation of energy recorded within previous shift by GMax from average energy recorded during eight previous shifts;

7. gdpuls: a deviation of a number of pulses recorded within previous shift by GMax from average number of pulses recorded during eight previous shifts;
8. ghazard: result of shift seismic hazard assessment in the mine working obtained by the seismoacoustic method based on registration coming from GMax only;
9. nbumps: the number of seismic bumps recorded within previous shift;
10. nbumps2: the number of seismic bumps (in energy range $[10^2, 10^3]$) registered within previous shift;
11. nbumps3: the number of seismic bumps (in energy range $[10^3, 10^4]$) registered within previous shift;
12. nbumps4: the number of seismic bumps (in energy range $[10^4, 10^5]$) registered within previous shift;
13. nbumps5: the number of seismic bumps (in energy range $[10^5, 10^6]$) registered within the last shift;
14. nbumps6: the number of seismic bumps (in energy range $[10^6, 10^7]$) registered within previous shift;
15. nbumps7: the number of seismic bumps (in energy range $[10^7, 10^8]$) registered within previous shift;
16. nbumps89: the number of seismic bumps (in energy range $[10^8, 10^{10}]$) registered within previous shift;
17. energy: total energy of seismic bumps registered within previous shift;
18. maxenergy: the maximum energy of the seismic bumps registered within previous shift;
19. class: the decision attribute - '1' means that high energy seismic bump occurred in the next shift ('hazardous state'), '0' means that no high energy seismic bumps occurred in the next shift ('non-hazardous state').

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.