# Codes created for the development of Hedorate.com and the creation of variables used to model property prices.

# Code 1: URL SCRAPPER

- Designed by Hedorate.com
- This code is designed for web scraping and data collection from a real estate website (duproprio.com) and utilizes R packages such as rvest, dplyr, and glue.

- Set a new working directory

```
setwd("/Users/simon-pierreboucher/Desktop/HEDORATE/webscrapping")
```

- Load the necessary libraries

```
library(rvest)
library(dplyr)
library(glue)
```

- Initialize an empty URL variable

```
url <- ""
```

- Create a loop to generate URLs for different pages of the website

```
for (i in 1:10) {
url[i] <-
paste0("https://duproprio.com/en/search/list?search=true&min_price=50000&subtype%5B0%5D=1&subtype%5B1%5D=6&subtype%5B2%5D=2&subtype%5B3%5D=7&subtype%5B4%5D=5&subtype%5B5%5D=4&rooms=1&bathrooms=1&lot_dimension_sq_feet=500~&living_space_sq_feet=100~&is_sold=1&parent=1&pageNumber=", i, "&sort=-published_at")
}
```

- Initialize empty variables to store links and dates

```
clinks <- ""
dates <- ""
```

- Loop through the generated URLs

```
for (i in 1:10) {
```

- Read the HTML content of the webpage

```
page <- read_html(url[i])
```

- Extract property links using CSS selectors

```
links <- page %>%
html_nodes(".search-results-listings-list__item-image-link") %>%
html_attr("href")
```

- Convert the links into a data frame

```
links <- as.data.frame(links)
```

- Extract sold dates using CSS selectors

```
date <- page %>%
html_nodes(".search-results-listings-list div.search-results-listings-list__item-description__item.search-results-listings-list__item-description__sold-in strong") %>%
html_text()
```

- Convert the dates into a data frame

```
date <- as.data.frame(date)
```

- Append the extracted links and dates to the respective variables

```
clinks <- rbind(clinks, links)
dates <- rbind(dates, date)
```

- Print a message indicating the progress

```
out <- paste0(i, "th property index page is scrapped")
print(out)
}
```

- Create a data frame from the collected links and dates

```
dub <- as.data.frame(cbind(clinks, dates))
```

# Code 2: HOUSING PAGE SCRAPPER

- Designed by Hedorate.com
- This code is intended for web scraping from a specific real estate website (duproprio.com) and organizing the scraped data into a structured format.

- Install and load the required library 'rvest'

```
install.packages("rvest")
library(rvest)
```

- Set the working directory to the specified path

```
setwd("/Users/simon-pierreboucher/Desktop/HEDORATE/webscrapping")
```

- Load a previously saved RData file named 'dup.RData'

```
load("dup.RData")
```

- Create an empty data frame named 'results' with predefined column names

```
results <- data.frame(
url = character(),
price = character(),
type = character(),
location = character(),
bedrooms = character(),
bathrooms = character(),
livingspace = character(),
lotdim = character(),
```

```
f1 = character(),
f2 = character(),
f3 = character(),
f4 = character(),
f5 = character(),
f6 = character(),
f7 = character(),
f8 = character(),
f9 = character(),
f10 = character(),
Fd1 = character(),
Fd2 = character(),
Fd3 = character(),
Fd4 = character(),
Fd5 = character(),
Fd6 = character(),
Fd7 = character(),
Fd8 = character(),
Fd9 = character(),
Fd10 = character(),
Fd11 = character(),
Fd12 = character(),
Fd13 = character(),
Fd14 = character(),
Fd15 = character(),
Fd16 = character(),
Fd17 = character(),
Fd18 = character(),
Fd19 = character(),
Fd20 = character(),
roomdim = character(),
solddate = character(),
stringsAsFactors = FALSE
)
```

- Loop to scrape data from a range of property URLs

```
for (i in 1:68941) { # You may adjust the range as needed
# Read the content of the webpage using the URL from the 'duproprio' dataset
page_content <- read_html(dup$links[i])
# Extract various property details using CSS selectors and store them in the 'results' data frame
results[i, 1] <- dup$links[i]
results[i, 2] <- page_content %>% html_node(".listing-sidebar > .listing-information > .listing-profile > .listing-address > .listing-price > .listing-price__amount") %>% html_text()
results[i, 3] <- page_content %>% html_node(".listing-sidebar > .listing-information > .listing-profile > .listing-address > .listing-location__title > a") %>% html_text()
results[i,4] <- page_content %>% html_node(".listing-sidebar > .listing-information > .listing-profile > .listing-address div.listing-location__address ") %>% html_text()
results[i,5] <- page_content %>% html_node(":nth-child(1) > .listing-main-characteristics__label > .listing-main-characteristics__number") %>% html_text()
results[i,6] <- page_content %>% html_node(".listing-main-characteristics__item--bathrooms > .listing-main-characteristics__label > .listing-main-characteristics__number") %>% html_text()
results[i,7] <- page_content %>% html_node(".listing-main-characteristics__item--living-space-area > .listing-main-characteristics__item-dimensions > .listing-main-characteristics__number") %>% html_text()
results[i,8] <- page_content %>% html_node(".listing-main-characteristics__item--lot-dimensions > .listing-main-
```

```
characteristics__item-dimensions > .listing-main-
characteristics__number") %>% html_text()
results[i,9] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(2)") %>% html_text()
results[i,10] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(3)") %>% html_text()
results[i,11] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(4)") %>% html_text()
results[i,12] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(5)") %>% html_text()
results[i,13] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(6)") %>% html_text()
results[i,14] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(7)") %>% html_text()
results[i,15] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(8)") %>% html_text()
results[i,16] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(9)") %>% html_text()
results[i,17] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(10)") %>% html_text()
results[i,18] <- page_content %>% html_node(".listing-list-
characteristics__viewport > :nth-child(11)") %>% html_text()
results[i,19] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(1)") %>% html_text()
results[i,20] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(2)") %>% html_text()
results[i,21] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(3)") %>% html_text()
results[i,22] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(4)") %>% html_text()
results[i,23] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(5)") %>% html_text()
results[i,24] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(6)") %>% html_text()
```

```
results[i,25] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(7)") %>% html_text()
results[i,26] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(8)") %>% html_text()
results[i,27] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(9)") %>% html_text()
results[i,28] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(10)") %>% html_text()
results[i,29] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(11)") %>% html_text()
results[i,30] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(12)") %>% html_text()
results[i,31] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(13)") %>% html_text()
results[i,32] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(14)") %>% html_text()
results[i,33] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(15)") %>% html_text()
results[i,34] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(16)") %>% html_text()
results[i,35] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(17)") %>% html_text()
results[i,36] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(18)") %>% html_text()
results[i,37] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(19)") %>% html_text()
results[i,38] <- page_content %>% html_node(".listing-complete-list-
characteristics__content > :nth-child(20)") %>% html_text()
results[i,39] <- page_content %>% html_node(".listing-rooms-details")
%>% html_text()
results[i,40] <- dup$date[i]
```

- Print a progress message

```
out <- paste0(i, "th property is scrapped, ", 68941 - i, " property remaining")
print(out)
}
```

- Store the scraped data in the 'bd' data frame

```
bd <- results
```

- Save the dataset 'bd' as an RData file named 'bd.RData'

```
save(bd, file = "bd.RData")
```

In summary, this R code loads a dataset of property URLs, scrapes detailed information from each property's webpage, and stores the scraped data in a structured data frame named 'results.' The code then saves this data frame as an RData file named 'bd.RData' for further analysis or use. Please note that you may need to adjust the URL range and CSS selectors based on the specific website structure and your data requirements.

# Code 3 : FEATURE ENGINEERING

- Designed by Hedorate.com

- Data preprocessing script for a dataset related to real estate or property listings. It's likely intended to clean and transform the data to make it more suitable for analysis.

- The code then proceeds with a series of data cleaning and transformation operations on the dataset "TEMP." The dataset "TEMP" is assumed to contain various columns, such as "price," "type," "location," "bedrooms," "bathrooms," "livingspace," and so on. The operations are as follows:

    - Several gsub functions are used to remove unwanted characters and spaces from the "price" column.

    - Similar gsub functions are used for the "type," "location," "bedrooms," "bathrooms," "livingspace," and "lotdim" columns.

    - The str_match function is used to extract specific patterns from the "location," "livingspace," and "lotdim" columns.

    - The "YEAR" column is extracted from the "solddate" column.

- A pattern is used to extract the "CONSTYEAR" (year of construction) from various columns like "f1," "f2," and so on.

- Similar patterns are used to extract various property features like "FACING," "FLOOR," "HEATING," "KITCHEN," "SHOWERANDBATH," "BASEMENT," "POOL," "GAR," and "LOCATION."

- Several binary variables (e.g., "FAC_BRICK," "FAC_STONE") are created based on the presence of specific features in the "FACING" column.

- Similar binary variables are created for "FLOOR," "HEATING," and other features.

```
setwd("/Users/simon-
pierreboucher/Desktop/HEDORATE/webscrapping")
load("bd.RData")

library(sjmisc)
library(stringr)

TEMP=bd
TEMP$price=gsub("\n","",TEMP$price)
TEMP$price=gsub(",","",TEMP$price)
TEMP$price=gsub(" ","",TEMP$price)
TEMP$price=gsub("SalepriceprovidedbythesellertoDuProprio.","",TE
MP$price)
TEMP$price=substr(TEMP$price,2,nchar(TEMP$price))
TEMP$type=gsub("\n ","",TEMP$type)
TEMP$type=gsub("\n ","",TEMP$type)
```

```
TEMP$location=gsub("\n \n ","%",TEMP$location)
TEMP$location=gsub("\n ","%%",TEMP$location)
TEMP$location=gsub("\n \n ","%%%",TEMP$location)
TEMP$TEMP1 <- str_match(TEMP$location, "%\s(.?)\s%%")
TEMP$CITY1<-TEMP$TEMP1[,2]
TEMP$TEMP2 <- str_match(TEMP$location, "%%\s(.?)\s%%%")
TEMP$CITY2<-TEMP$TEMP2[,2]

TEMP$bedrooms=gsub("\n","",TEMP$bedrooms)
TEMP$bedrooms=gsub(" ","",TEMP$bedrooms)
TEMP$bedrooms=as.numeric(TEMP$bedrooms)

TEMP$bathrooms=gsub("\n","",TEMP$bathrooms)
TEMP$bathrooms=gsub(" ","",TEMP$bathrooms)
TEMP$bathrooms=as.numeric(TEMP$bathrooms)

TEMP$livingspace=gsub("\n ","%",TEMP$livingspace)
TEMP$livingspace=gsub(" ft²","%%",TEMP$livingspace)
TEMP$livingspace=gsub(" ft","%%%",TEMP$livingspace)
TEMP$LA <- str_match(TEMP$livingspace, "%\s(.?)\s*%%")
TEMP$LA=gsub(" ","",TEMP$LA)
TEMP$LA=gsub(",","",TEMP$LA)
TEMP$LA=TEMP$LA[,2]
TEMP$LA_TEMP1 <- str_match(TEMP$livingspace, "%\s(.?)\sx")
TEMP$LA_TEMP2 <- str_match(TEMP$livingspace, "x\s(.?)\s%%%")
TEMP$LA_TEMP <-
as.numeric(TEMP$LA_TEMP1[,2])*as.numeric(TEMP$LA_TEMP2[,
2])
TEMP$LA<- ifelse(is.na(TEMP$LA_TEMP),TEMP$LA
,TEMP$LA_TEMP)
TEMP$LA=as.numeric(TEMP$LA)
```

```
TEMP$lotdim=gsub("\n ","%",TEMP$lotdim)
TEMP$lotdim=gsub(" ft²","%%",TEMP$lotdim)
TEMP$lotdim=gsub(" ft","%%%",TEMP$lotdim)
TEMP$LD <- str_match(TEMP$lotdim, "%\s(.?)\s%%")
TEMP$LD=gsub(" ","",TEMP$LD)
TEMP$LD=gsub(",","",TEMP$LD)
TEMP$LD=TEMP$LD[,2]
TEMP$LD_TEMP1 <- str_match(TEMP$lotdim, "%\s(.?)\sx")
TEMP$LD_TEMP2 <- str_match(TEMP$lotdim, "x\s(.?)\s%%%")
TEMP$LD_TEMP <-
as.numeric(TEMP$LD_TEMP1[,2])as.numeric(TEMP$LD_TEMP2[,2
])
TEMP$LD<- ifelse(is.na(TEMP$LD_TEMP),TEMP$LD
,TEMP$LD_TEMP)
TEMP$LD=as.numeric(TEMP$LD)

TEMP$YEAR=substr(TEMP$solddate,nchar(TEMP$solddate)-
3,nchar(TEMP$solddate))
TEMP$YEAR_temp=as.numeric(TEMP$YEAR)
pattern<-"Year of construction\n"
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f1)==1,
TEMP$f1, "")
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f2)==1,
TEMP$f2, TEMP$CONSTYEAR)
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f3)==1,
TEMP$f3, TEMP$CONSTYEAR)
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f4)==1,
TEMP$f4, TEMP$CONSTYEAR)
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f5)==1,
TEMP$f5, TEMP$CONSTYEAR)
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f6)==1,
TEMP$f6, TEMP$CONSTYEAR)
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f7)==1,
TEMP$f7, TEMP$CONSTYEAR)
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f8)==1,
TEMP$f8, TEMP$CONSTYEAR)
```

```
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f9)==1,
TEMP$f9, TEMP$CONSTYEAR)
TEMP$CONSTYEAR<- ifelse(grepl(pattern,TEMP$f10)==1,
TEMP$f10, TEMP$CONSTYEAR)
TEMP$CONSTYEAR=gsub("\n Year of construction\n \n
","",TEMP$CONSTYEAR)
TEMP$CONSTYEAR=gsub("\n ","",TEMP$CONSTYEAR)
TEMP$CONSTYEAR=as.numeric(TEMP$CONSTYEAR)
TEMP$AGE=ifelse(grepl("New",TEMP$type)==1,0
,TEMP$YEAR_temp-TEMP$CONSTYEAR)
```

```
pattern<-"External facing:\n"
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd1)==1, TEMP$Fd1,
"")
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd2)==1, TEMP$Fd2,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd3)==1, TEMP$Fd3,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd4)==1, TEMP$Fd4,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd5)==1, TEMP$Fd5,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd6)==1, TEMP$Fd6,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd7)==1, TEMP$Fd7,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd8)==1, TEMP$Fd8,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd9)==1, TEMP$Fd9,
TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd10)==1,
TEMP$Fd10, TEMP$FACING)
```

```
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd11)==1,
TEMP$Fd11, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd12)==1,
TEMP$Fd12, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd13)==1,
TEMP$Fd13, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd14)==1,
TEMP$Fd14, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd15)==1,
TEMP$Fd15, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd16)==1,
TEMP$Fd16, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd17)==1,
TEMP$Fd17, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd18)==1,
TEMP$Fd18, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd19)==1,
TEMP$Fd19, TEMP$FACING)
TEMP$FACING<- ifelse(grepl(pattern,TEMP$Fd20)==1,
TEMP$Fd20, TEMP$FACING)


pattern<-"Floor coverings:\n"
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd1)==1, TEMP$Fd1,
"")
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd2)==1, TEMP$Fd2,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd3)==1, TEMP$Fd3,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd4)==1, TEMP$Fd4,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd5)==1, TEMP$Fd5,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd6)==1, TEMP$Fd6,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd7)==1, TEMP$Fd7,
TEMP$FLOOR)
```

```
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd8)==1, TEMP$Fd8,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd9)==1, TEMP$Fd9,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd10)==1, TEMP$Fd10,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd11)==1, TEMP$Fd11,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd12)==1, TEMP$Fd12,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd13)==1, TEMP$Fd13,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd14)==1, TEMP$Fd14,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd15)==1, TEMP$Fd15,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd16)==1, TEMP$Fd16,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd17)==1, TEMP$Fd17,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd18)==1, TEMP$Fd18,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd19)==1, TEMP$Fd19,
TEMP$FLOOR)
TEMP$FLOOR<- ifelse(grepl(pattern,TEMP$Fd20)==1, TEMP$Fd20,
TEMP$FLOOR)


pattern<-"Heating source:\n"
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd1)==1,
TEMP$Fd1, "")
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd2)==1,
TEMP$Fd2, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd3)==1,
TEMP$Fd3, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd4)==1,
TEMP$Fd4, TEMP$HEATING)
```

```
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd5)==1,
TEMP$Fd5, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd6)==1,
TEMP$Fd6, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd7)==1,
TEMP$Fd7, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd8)==1,
TEMP$Fd8, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd9)==1,
TEMP$Fd9, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd10)==1,
TEMP$Fd10, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd11)==1,
TEMP$Fd11, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd12)==1,
TEMP$Fd12, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd13)==1,
TEMP$Fd13, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd14)==1,
TEMP$Fd14, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd15)==1,
TEMP$Fd15, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd16)==1,
TEMP$Fd16, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd17)==1,
TEMP$Fd17, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd18)==1,
TEMP$Fd18, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd19)==1,
TEMP$Fd19, TEMP$HEATING)
TEMP$HEATING<- ifelse(grepl(pattern,TEMP$Fd20)==1,
TEMP$Fd20, TEMP$HEATING)
```

```
pattern<-"Kitchen:\n"
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd1)==1, TEMP$Fd1,
"")
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd2)==1, TEMP$Fd2,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd3)==1, TEMP$Fd3,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd4)==1, TEMP$Fd4,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd5)==1, TEMP$Fd5,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd6)==1, TEMP$Fd6,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd7)==1, TEMP$Fd7,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd8)==1, TEMP$Fd8,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd9)==1, TEMP$Fd9,
TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd10)==1,
TEMP$Fd10, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd11)==1,
TEMP$Fd11, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd12)==1,
TEMP$Fd12, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd13)==1,
TEMP$Fd13, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd14)==1,
TEMP$Fd14, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd15)==1,
TEMP$Fd15, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd16)==1,
TEMP$Fd16, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd17)==1,
TEMP$Fd17, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd18)==1,
TEMP$Fd18, TEMP$KITCHEN)
```

```
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd19)==1,
TEMP$Fd19, TEMP$KITCHEN)
TEMP$KITCHEN<- ifelse(grepl(pattern,TEMP$Fd20)==1,
TEMP$Fd20, TEMP$KITCHEN)


pattern<-"Bathroom:\n"
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd1)==1,
TEMP$Fd1, "")
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd2)==1,
TEMP$Fd2, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd3)==1,
TEMP$Fd3, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd4)==1,
TEMP$Fd4, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd5)==1,
TEMP$Fd5, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd6)==1,
TEMP$Fd6, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd7)==1,
TEMP$Fd7, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd8)==1,
TEMP$Fd8, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd9)==1,
TEMP$Fd9, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd10)==1,
TEMP$Fd10, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd11)==1,
TEMP$Fd11, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd12)==1,
TEMP$Fd12, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd13)==1,
TEMP$Fd13, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd14)==1,
TEMP$Fd14, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd15)==1,
TEMP$Fd15, TEMP$SHOWERANDBATH)
```

```
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd16)==1,
TEMP$Fd16, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd17)==1,
TEMP$Fd17, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd18)==1,
TEMP$Fd18, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd19)==1,
TEMP$Fd19, TEMP$SHOWERANDBATH)
TEMP$SHOWERANDBATH<- ifelse(grepl(pattern,TEMP$Fd20)==1,
TEMP$Fd20, TEMP$SHOWERANDBATH)


pattern<-"Basement:\n"
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd1)==1,
TEMP$Fd1, "")
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd2)==1,
TEMP$Fd2, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd3)==1,
TEMP$Fd3, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd4)==1,
TEMP$Fd4, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd5)==1,
TEMP$Fd5, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd6)==1,
TEMP$Fd6, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd7)==1,
TEMP$Fd7, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd8)==1,
TEMP$Fd8, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd9)==1,
TEMP$Fd9, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd10)==1,
TEMP$Fd10, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd11)==1,
TEMP$Fd11, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd12)==1,
TEMP$Fd12, TEMP$BASEMENT)
```

```
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd13)==1,
TEMP$Fd13, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd14)==1,
TEMP$Fd14, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd15)==1,
TEMP$Fd15, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd16)==1,
TEMP$Fd16, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd17)==1,
TEMP$Fd17, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd18)==1,
TEMP$Fd18, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd19)==1,
TEMP$Fd19, TEMP$BASEMENT)
TEMP$BASEMENT<- ifelse(grepl(pattern,TEMP$Fd20)==1,
TEMP$Fd20, TEMP$BASEMENT)


pattern<-"Pool:\n"
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd1)==1, TEMP$Fd1, "")
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd2)==1, TEMP$Fd2,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd3)==1, TEMP$Fd3,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd4)==1, TEMP$Fd4,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd5)==1, TEMP$Fd5,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd6)==1, TEMP$Fd6,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd7)==1, TEMP$Fd7,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd8)==1, TEMP$Fd8,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd9)==1, TEMP$Fd9,
TEMP$POOL)
```

```
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd10)==1, TEMP$Fd10,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd11)==1, TEMP$Fd11,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd12)==1, TEMP$Fd12,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd13)==1, TEMP$Fd13,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd14)==1, TEMP$Fd14,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd15)==1, TEMP$Fd15,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd16)==1, TEMP$Fd16,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd17)==1, TEMP$Fd17,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd18)==1, TEMP$Fd18,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd19)==1, TEMP$Fd19,
TEMP$POOL)
TEMP$POOL<- ifelse(grepl(pattern,TEMP$Fd20)==1, TEMP$Fd20,
TEMP$POOL)


pattern<-"Garage:\n"
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd1)==1, TEMP$Fd1, "")
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd2)==1, TEMP$Fd2,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd3)==1, TEMP$Fd3,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd4)==1, TEMP$Fd4,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd5)==1, TEMP$Fd5,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd6)==1, TEMP$Fd6,
TEMP$GAR)
```

```
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd7)==1, TEMP$Fd7,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd8)==1, TEMP$Fd8,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd9)==1, TEMP$Fd9,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd10)==1, TEMP$Fd10,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd11)==1, TEMP$Fd11,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd12)==1, TEMP$Fd12,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd13)==1, TEMP$Fd13,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd14)==1, TEMP$Fd14,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd15)==1, TEMP$Fd15,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd16)==1, TEMP$Fd16,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd17)==1, TEMP$Fd17,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd18)==1, TEMP$Fd18,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd19)==1, TEMP$Fd19,
TEMP$GAR)
TEMP$GAR<- ifelse(grepl(pattern,TEMP$Fd20)==1, TEMP$Fd20,
TEMP$GAR)


pattern<-"Location:\n"
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd1)==1,
TEMP$Fd1, "")
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd2)==1,
TEMP$Fd2, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd3)==1,
TEMP$Fd3, TEMP$LOCATION)
```

```
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd4)==1,
TEMP$Fd4, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd5)==1,
TEMP$Fd5, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd6)==1,
TEMP$Fd6, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd7)==1,
TEMP$Fd7, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd8)==1,
TEMP$Fd8, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd9)==1,
TEMP$Fd9, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd10)==1,
TEMP$Fd10, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd11)==1,
TEMP$Fd11, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd12)==1,
TEMP$Fd12, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd13)==1,
TEMP$Fd13, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd14)==1,
TEMP$Fd14, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd15)==1,
TEMP$Fd15, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd16)==1,
TEMP$Fd16, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd17)==1,
TEMP$Fd17, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd18)==1,
TEMP$Fd18, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd19)==1,
TEMP$Fd19, TEMP$LOCATION)
TEMP$LOCATION<- ifelse(grepl(pattern,TEMP$Fd20)==1,
TEMP$Fd20, TEMP$LOCATION)
```

```
TEMP$FAC_BRICK<- ifelse(grepl("Brick\n",TEMP$FACING)==1,
1, 0)
TEMP$FAC_STONE<- ifelse(grepl("Stone\n",TEMP$FACING)==1,
1, 0)
TEMP$FAC_VINYL<- ifelse(grepl("Vinyl
Siding\n",TEMP$FACING)==1, 1, 0)
TEMP$FAC_WOOD<- ifelse(grepl("Wood\n",TEMP$FACING)==1,
1, 0)
TEMP$FAC_ALLUMINUIUM<- ifelse(grepl("Aluminium
Siding\n",TEMP$FACING)==1, 1, 0)
TEMP$FLOOR_CERAMIC<-
ifelse(grepl("Ceramic\n",TEMP$FLOOR)==1, 1, 0)
TEMP$FLOOR_HARDWOOD<-
ifelse(grepl("Hardwood\n",TEMP$FLOOR)==1, 1, 0)
TEMP$FLOOR_LAMINATE<-
ifelse(grepl("Laminate\n",TEMP$FLOOR)==1, 1, 0)
TEMP$HEATING_FORCEDAIR<- ifelse(grepl("Forced
air\n",TEMP$HEATING)==1, 1, 0)
TEMP$HEATING_NATURALGAS<- ifelse(grepl("Natural
gas\n",TEMP$HEATING)==1, 1, 0)
TEMP$HEATING_ELECTRICITY<-
ifelse(grepl("Electric\n",TEMP$HEATING)==1, 1, 0)
TEMP$BASEMENT_FINISH<- ifelse(grepl("Totally
finished\n",TEMP$BASEMENT)==1, 1, 0)
TEMP$POOL_ABOVE<- ifelse(grepl("Above
ground\n",TEMP$POOL)==1, 1, 0)
TEMP$POOL_IN<- ifelse(grepl("Inground\n",TEMP$POOL)==1, 1,
0)
TEMP$GAR_SINGLE<- ifelse(grepl("Single\n",TEMP$GAR)==1, 1,
0)
TEMP$GAR_DOUBLE<- ifelse(grepl("Double\n",TEMP$GAR)==1,
1, 0)
TEMP$LOCATION_NOBACK<- ifelse(grepl("No backyard
neighbors\n",TEMP$LOCATION)==1, 1, 0)
TEMP$LOCATION_RESIDENTIAL<- ifelse(grepl("Residential
area\n",TEMP$LOCATION)==1, 1, 0)
```

```
TEMP$LOCATION_PARK<- ifelse(grepl("Near
park\n",TEMP$LOCATION)==1, 1, 0)
TEMP$LOCATION_PUBLICT<- ifelse(grepl("Public
transportation\n",TEMP$LOCATION)==1, 1, 0)
TEMP$LOCATION_HIGHWAY<- ifelse(grepl("Highway
access\n",TEMP$LOCATION)==1, 1, 0)
TEMP$type=ifelse(TEMP$type=="New 2 Storey", "2
Storey",TEMP$type)
TEMP$type=ifelse(TEMP$type=="New Bungalow",
"Bungalow",TEMP$type)
TEMP$type=ifelse(TEMP$type=="New Semi-detached", "Semi-
detached",TEMP$type)
TEMP$type=ifelse(TEMP$type=="New Townhouse",
"Townhouse",TEMP$type)
TEMP$type=ifelse(TEMP$type=="New 1 1/2 Storey", "1 1/2
Storey",TEMP$type)
```

```
DB_REGRESS=as.data.frame(cbind(TEMP$price,TEMP$type,TEMP$
CITY1,TEMP$bedrooms,TEMP$bathrooms,TEMP$LA,TEMP$LD,T
EMP$AGE,TEMP$YEAR,TEMP$POOL_ABOVE,TEMP$POOL_IN,
TEMP$FAC_BRICK,TEMP$FAC_STONE,TEMP$FLOOR_CERAM
IC,TEMP$FLOOR_HARDWOOD,TEMP$GAR_DOUBLE,TEMP$G
AR_SINGLE,TEMP$HEATING_ELECTRICITY,TEMP$HEATING_
NATURALGAS,TEMP$BASEMENT_FINISH,TEMP$LOCATION_
HIGHWAY,TEMP$LOCATION_NOBACK,TEMP$LOCATION_RE
SIDENTIAL,TEMP$LOCATION_PUBLICT))
names(DB_REGRESS)[1]<-"PRICE"
names(DB_REGRESS)[2]<-"TYPE"
names(DB_REGRESS)[3]<-"CITY"
names(DB_REGRESS)[4]<-"BED"
names(DB_REGRESS)[5]<-"BATH"
names(DB_REGRESS)[6]<-"LA"
names(DB_REGRESS)[7]<-"LD"
names(DB_REGRESS)[8]<-"AGE"
names(DB_REGRESS)[9]<-"YOS"
names(DB_REGRESS)[10]<-"POOL_ABOVE"
names(DB_REGRESS)[11]<-"POOL_IN"
names(DB_REGRESS)[12]<-"BRICK"
names(DB_REGRESS)[13]<-"STONE"
names(DB_REGRESS)[14]<-"CERAMIC"
names(DB_REGRESS)[15]<-"HARDWOOD"
names(DB_REGRESS)[16]<-"GAR_DOUBLE"
names(DB_REGRESS)[17]<-"GAR_SINGLE"
names(DB_REGRESS)[18]<-"ELECTRICITY"
names(DB_REGRESS)[19]<-"NATURALGAS"
names(DB_REGRESS)[20]<-"BASEMENT_FINISH"
names(DB_REGRESS)[21]<-"HIGHWAY"
names(DB_REGRESS)[22]<-"NOBACK"
names(DB_REGRESS)[23]<-"RESIDENTIAL"
names(DB_REGRESS)[24]<-"PUBLICT"
```

```
DB_REGRESS$PRICE=as.numeric(DB_REGRESS$PRICE)
DB_REGRESS$BED=as.numeric(DB_REGRESS$BED)
DB_REGRESS$BATH=as.numeric(DB_REGRESS$BATH)
DB_REGRESS$LA=as.numeric(DB_REGRESS$LA)
DB_REGRESS$LD=as.numeric(DB_REGRESS$LD)
DB_REGRESS$AGE=as.numeric(DB_REGRESS$AGE)
DB_REGRESS$TYPE=as.factor(DB_REGRESS$TYPE)
DB_REGRESS$YOS=as.factor(DB_REGRESS$YOS)
DB_REGRESS$CITY=as.factor(DB_REGRESS$CITY)
YY=na.omit(DB_REGRESS)

save(YY,file = "BD_HEDONIC.Rdata")


HOUSING_MODEL<-lm(PRICE ~ BED + BATH + LA + LD + AGE
+ POOL_ABOVE + POOL_IN + GAR_DOUBLE + GAR_SINGLE +
BASEMENT_FINISH + BRICK + STONE + CERAMIC +
HARDWOOD + HIGHWAY + NOBACK + RESIDENTIAL +
PUBLICT + TYPE + YOS + CITY ,data = YY)
HEDONIC_COEF=as.data.frame(HOUSING_MODEL$coefficients)
HEDONIC_COEF$NAME=row.names(HEDONIC_COEF)
write.table(HEDONIC_COEF,file = "HEDONIC_COEF.txt")
```

The code performed data cleaning and feature engineering on a real estate dataset. It creates new variables and cleans existing ones to prepare the data for analysis or modeling.