homework

Yuwei Sun

2017/2/5

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.3.2
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
## Warning: package 'tidyr' was built under R version 3.3.2
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

2016 bridge data of Wisconsin

```
link = "https://www.fhwa.dot.gov/bridge/nbi/2016/delimited/WI16.txt"
WI16 = read_csv(link)
```

```
## Parsed with column specification:
## cols(
##   .default = col_character(),
##   STATE_CODE_001 = col_integer(),
##   RECORD_TYPE_005A = col_integer(),
##   ROUTE_PREFIX_005B = col_integer(),
##   SERVICE_LEVEL_005C = col_integer(),
##   DIRECTION_005E = col_integer(),
##   PLACE_CODE_004 = col_integer(),
##   MIN_VERT_CLR_010 = col_double(),
##   KILOPOINT_011 = col_double(),
##   BASE_HWY_NETWORK_012 = col_integer(),
##   SUBROUTE_NO_013B = col_integer(),
##   DETOUR_KILOS_019 = col_integer(),
##   TOLL_020 = col_integer(),
##   YEAR_BUILT_027 = col_integer(),
##   TRAFFIC_LANES_ON_028A = col_integer(),
##   TRAFFIC_LANES_UND_028B = col_integer(),
##   ADT_029 = col_integer(),
```

```
##   YEAR_ADT_030 = col_integer(),
##   APPR_WIDTH_MT_032 = col_double(),
##   MEDIAN_CODE_033 = col_integer(),
##   DEGREES_SKEW_034 = col_integer()
##   # ... with 54 more columns
## )
```

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

```
## Warning: 29 parsing failures.
##  row                     col   expected          actual
## 3058 OTHR_STATE_STRUC_NO_099 an integer 22122011000B010
## 3061 OTHR_STATE_STRUC_NO_099 an integer 22122021000B020
## 3065 OTHR_STATE_STRUC_NO_099 an integer 36136031000B010
## 3069 OTHR_STATE_STRUC_NO_099 an integer 36136051000B010
## 3280 OTHR_STATE_STRUC_NO_099 an integer 36136011000B010
## .... ...................... .......... ..............
## See problems(...) for more details.
```

Find something about the improvement cost.

```
wi=transmute(WI16,year=YEAR_BUILT_027,lat=LAT_016,lon=LONG_017,adt=ADT_029,costb=BRIDGE_IMP_COST_094,
           costr=ROADWAY_IMP_COST_095,costt=TOTAL_IMP_COST_096)
#Many cells are zero,some are NA...
wi=filter(wi,costb>0,costr>0,costt>2500)
```

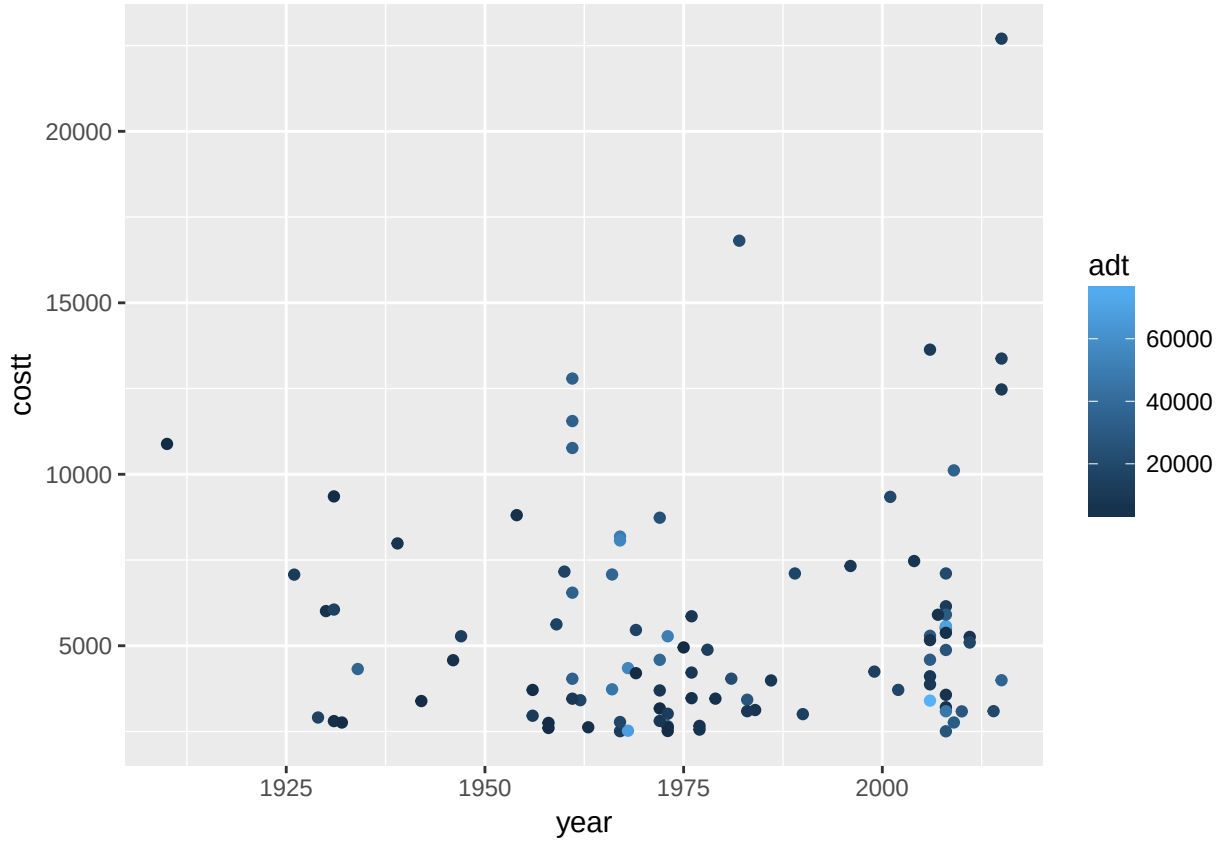
Group by year built,find the mean total improvement cost.

```
wi%>%
  group_by(year)%>%
  summarise(mean=mean(costt,na.rm=T))
```

```
## # A tibble: 50 <U+00D7> 2
##     year  mean
##    <int> <dbl>
## 1   1910 10885
## 2   1926  7074
## 3   1929  2908
## 4   1930  6011
## 5   1931  6071
## 6   1932  2762
## 7   1934  4322
## 8   1939  7984
## 9   1942  3388
## 10  1946  4578
## # ... with 40 more rows
```

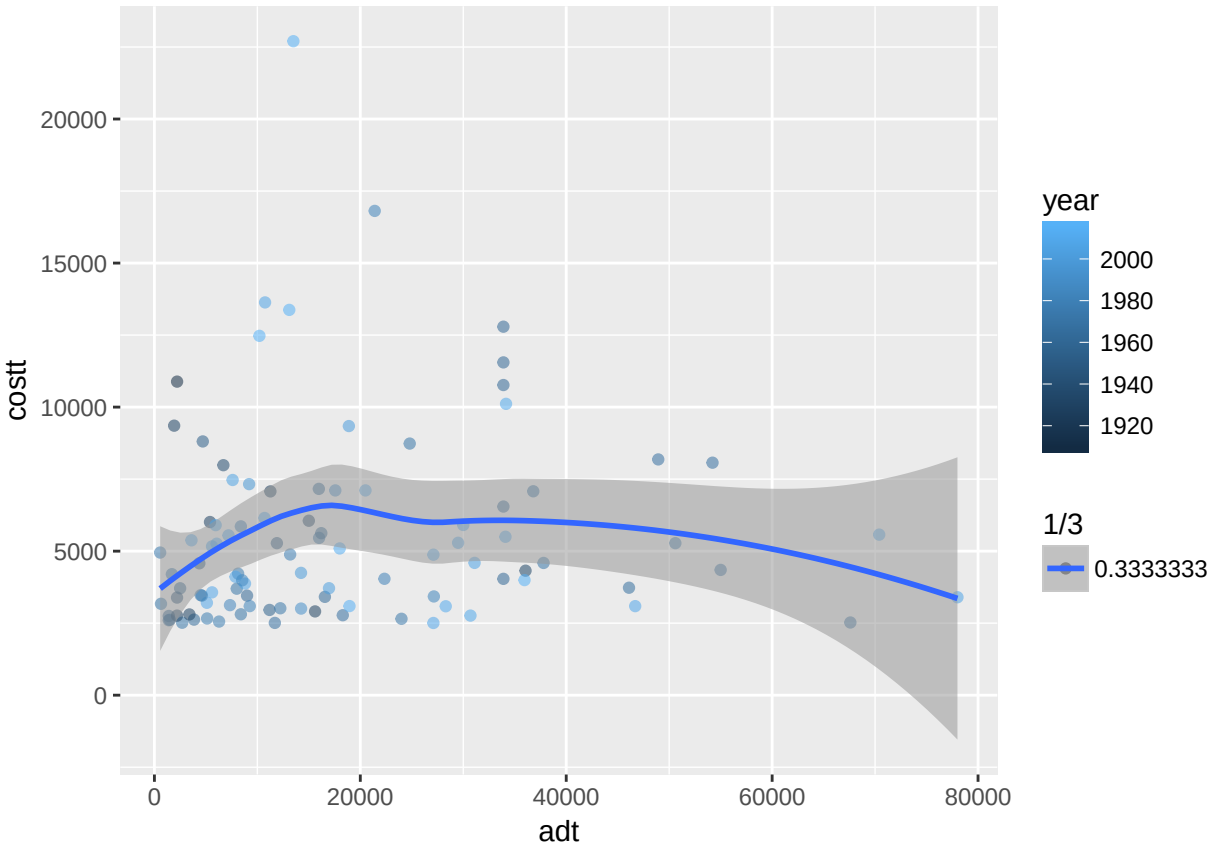
Find if the total improvement cost has some relationship with other variables,like year built,average daily traffic etc.

```
ggplot(wi)+geom_point(mapping=aes(x=year,y=costt,color=adt))
```




```
ggplot(wi,mapping=aes(x=adt,y=costt,alpha=1/3))+
  geom_point(mapping=aes(col=year))+
  geom_smooth()
```

```
## `geom_smooth()` using method = 'loess'
```


20000
15000
10000
5000
0
costt
0
20000
40000
60000
80000
adt
year
2000
1980
1960
1940
1920
1/3
0.3333333