---

title: "Data607\_Project2\_Rajan"

output:

pdf\_document: default

html\_document: default

---

```{r}

require(dplyr)

require(tidyr)

require(magrittr)

require(stringr)

require(zoo)

require(ggplot2)

require(ggthemes)

require(extrafont)

##Fix poor formating from the file

AeroCost = read.csv(file="AeroCost.csv",header=TRUE,sep=",",na.strings = c("","NA"))

head(AeroCost)

Type

<fctr>

Model

<fctr>

X1941

<fctr>

X1942

<fctr>

X1943

<fctr>

X1944

<fctr>

X1945

<fctr>

1 Very Heavy Bombers NA NA NA NA NA NA

2 NA B-29 NA 897,730 NA 605,360 509,465

3 Heavy Bombers NA NA NA NA NA NA

4 NA B-17 301,221 258,949 NA 204,370 187,742

5 NA B-24 379,162 304,391 NA 215,516 NA

6 NA B-32 NA 790,433 NA 790,433 NA

##Tidy data done with TidyR,

AeroCost$Type=na.locf(AeroCost$Type)

AeroCost$Type = str\_trim(AeroCost$Type)

AeroCost$Model = str\_trim(AeroCost$Model)

AeroCost$Type=as.factor(AeroCost$Type)

AeroCost$Model= as.factor(AeroCost$Model)

AeroCost = gather(AeroCost,Year,Cost,X1941:X1945)

AeroCost$Model[AeroCost$Model==""] <- NA

AeroCost$Model= as.factor(AeroCost$Model)

AeroCost$Year=str\_replace(AeroCost$Year,"X","")

AeroCost$Year=as.factor(AeroCost$Year)

AeroCost$Cost=str\_replace(AeroCost$Cost,",","")

AeroCost$Cost=as.numeric(AeroCost$Cost)

AeroCost = AeroCost[complete.cases(AeroCost$Model),]

write.csv(AeroCost,"Aircraftclean.csv", row.names=FALSE)

head(AeroCost)

|  | **Type**  <fctr> | **Model**  <fctr> | **Year**  <fctr> | **Cost**  <dbl> |
| --- | --- | --- | --- | --- |
| 2 | Very Heavy Bombers | B-29 | 1941 | *NA* |
| 4 | Heavy Bombers | B-17 | 1941 | 301221 |
| 5 | Heavy Bombers | B-24 | 1941 | 379162 |
| 6 | Heavy Bombers | B-32 | 1941 | *NA* |
| 8 | Medium Bombers | B-25 | 1941 | 180031 |
| 9 | Medium Bombers | B-26 | 1941 | 261062 |

6 rows

##Plot to study cost of Transports.

library(dplyr)

library(tidyr)

target=c("C-43", "C-45", "C-46", "C-47", "C-53", "C-54", "C-60", "C-61", "UC-64", "C-69", "C-74", "C-78", "C-82", "C-87")

Transports=dplyr::filter(c,Model %in%target)

g=ggplot(data=Transports,aes(x=Year,y=Cost\*0.001,group = Model, color = Model))

g=g+geom\_line(stat="identity",size=1.6)

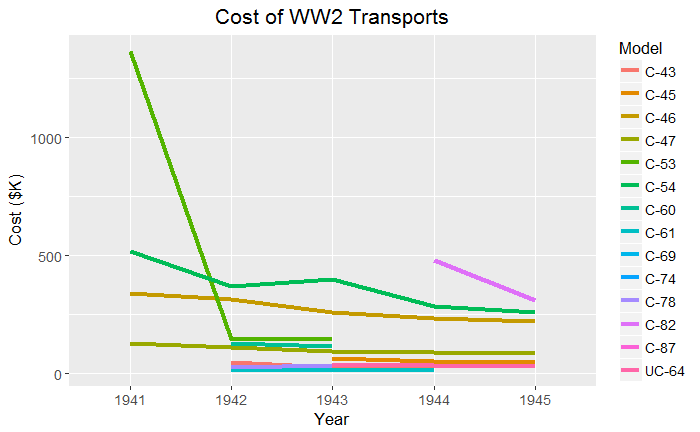
g=g+ggtitle("Cost of WW2 Transports")

g=g+ylab("Cost ($K)")+xlab("Year")

g=g+theme\_get()

g=g+theme(plot.title = element\_text(hjust = 0.5),text=element\_text(size=13, family="Times"))

g



```