Blood Storage

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.5.2     v tibble    3.3.0
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.1.0
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
```

```
blood <- read.csv("BloodStorage.csv")

names(blood)
```

```
 [1] "RBC.Age.Group"    "Median.RBC.Age"   "Age"              "AA"
 [5] "FamHx"            "PVol"             "TVol"             "T.Stage"
 [9] "bGS"              "BN."              "OrganConfined"    "PreopPSA"
[13] "PreopTherapy"     "Units"            "sGS"              "AnyAdjTherapy"
[17] "AdjRadTherapy"    "Recurrence"       "Censor"           "TimeToRecurrence"
```

```
blood$`RBC.Age.Group` <- factor(blood$`RBC.Age.Group`,
                                levels = c(1, 2, 3),
                                labels = c("Younger", "Middle", "Older"))
anova_result <- aov(TimeToRecurrence ~ RBC.Age.Group, data = blood)
summary(anova_result)
```

```
                Df Sum Sq Mean Sq F value Pr(>F)
RBC.Age.Group    2    771   385.5   0.471  0.625
Residuals      312 255329   818.4
1 observation deleted due to missingness
```

```
[1] 4
```

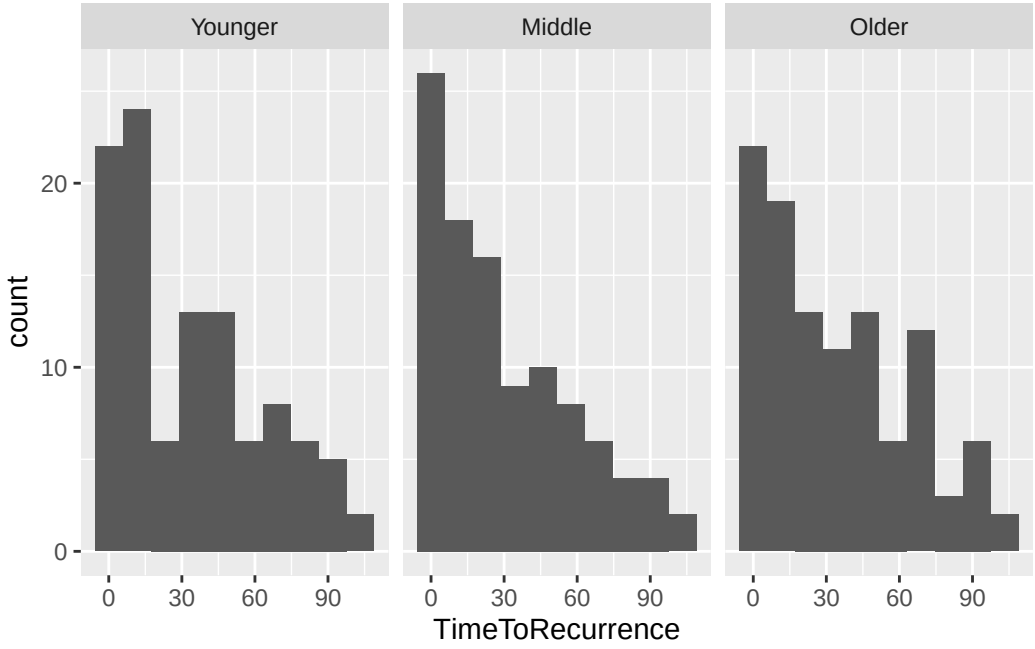
Assumptions 1) The samples are independent.

This assumption is reasonably met by the nature of the data. Each sample represents and independent individual patient. Any condtions for one patient would not impact the conditions of another patient.

2) The variance is the same within all groups

```
library(ggplot2)
ggplot(data = blood, aes(x = TimeToRecurrence)) +
  geom_histogram(bins = 10) +
  facet_grid(~ RBC.Age.Group)
```

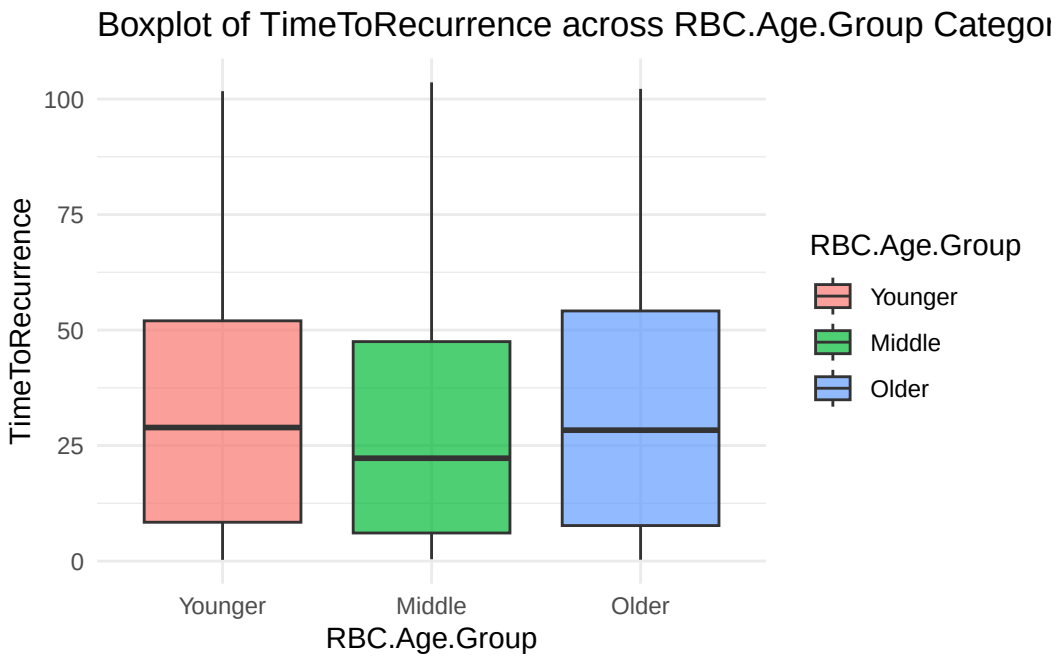
```
Warning: Removed 1 row containing non-finite outside the scale range
(`stat_bin()`).
```



This assumption is not met. Outcomes between groups of Blood Age (Younger, Middle, Older) are not normally distributed. Rather, they are right skewed in each group.

3) Outcomes within groups are normally distributed

```
Warning: Removed 1 row containing non-finite outside the scale range
(`stat_boxplot()`).
```



This assumption of anova is met. The variance across groups is nearly identical.