Academic integrity statement

You must sign this (typing in your details is acceptable) and include it with each piece of work you submit.

I am aware that the University defines plagiarism as presenting someone else's work, in whole or in part, as your own. Work means any intellectual output, and typically includes text, data, images, sound or performance.

I promise that in the attached submission I have not presented anyone else's work, in whole or in part, as my own and I have not colluded with others in the preparation of this work. Where I have taken advantage of the work of others, I have given full acknowledgement. I have not resubmitted my own work or part thereof without specific written permission to do so from the University staff concerned when any of this work has been or is being submitted for marks or credits even if in a different module or for a different qualification or completed prior to entry to the University. I have read and understood the University's published rules on plagiarism and also any more detailed rules specified at School or module level. I know that if I commit plagiarism I can be expelled from the University and that it is my responsibility to be aware of the University's regulations on plagiarism and their importance.

I re-confirm my consent to the University copying and distributing any or all of my work in any form and using third parties (who may be based outside the EU/EEA) to monitor breaches of regulations, to verify whether my work contains plagiarised material, and for quality assurance purposes.

I confirm that I have declared all mitigating circumstances that may be relevant to the assessment of this piece of work and that I wish to have taken into account. I am aware of the University's policy on mitigation and the School's procedures for the submission of statements and evidence of mitigation. I am aware of the penalties imposed for the late submission of coursework.

Name	Adnan Malik
Student ID	201538993

Adnan Malik
201538993

Effects of "Motives behind murder" on the Career-Timeline of a Serial killer.

Introduction

A serial killer in definition is a person who is responsible for a series of murders committed for a considerable time-period (Career-Timeline). Career-Timeline includes a starting and ending point of career (i.e., Age at First and Last kill) and the duration between the start and end point for which the killer was active.

The intent is to present an in-detail analysis

- To know how Motive behind murder affects the starting point of a serial killer's career. This will allow us to know whether a difference in motive can lead to some killers start younger than others.
- Also, to confirm the findings of a previously done research which suggested that on an average the killer's "Age at First kill" is 27.

The Motives behind murder considered for this analysis are the following three:

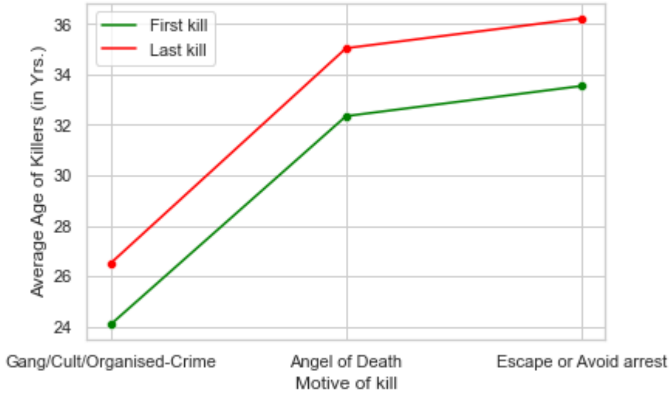
1. Gang, cult or organised crime:
   Killers associated with gang, cult or any form of organised crime.
2. Angel of Death:
   People who work as a caretaker (e.g., nurses) and intentionally murders those who are under their care.
3. Escape or avoid arrest:
   Killers who kill in order to avoid arrest or escape from custody.

Results

The analysis is done after cleaning the data from Null values, special values and outliers, which resulted in a loss of 10% (approx.) data. Now, to understand our data of Age at first kill by Motive more properly look at the table below.

	Gang, cult or Organised crime (n=93)	Angel of Death (n=23)	Killed to escape or avoid arrest (n=22)
Age at First Kill	23 (5)	30 (11)	31 (12)
Age at Last Kill	26 (8)	33 (13)	32 (13)
Career Duration	1 (3)	1 (1)	0 (1)

Table 1. Median (Interquartile range) of data for Age at First kill (in Yrs.), Age at Last kill (in Yrs.) and Career duration (in Yrs.) for killer with different "Motives of Murder".



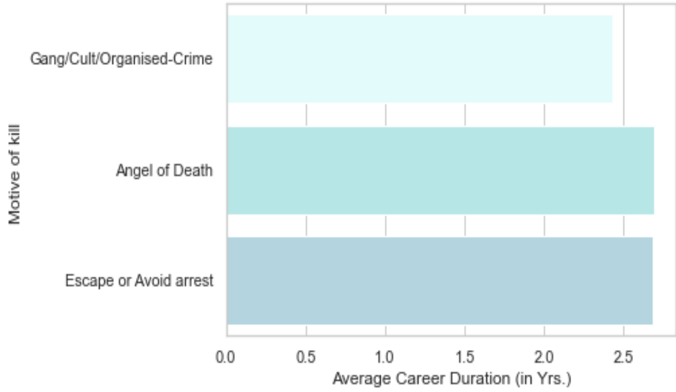


Figure 1. The line-plot (left) shows a relationship between Avg. age of killers (in Yrs.) and the motive. The bar-plot (right) shows the Avg. career duration (in Yrs.) for each Motive of murder

The plots above suggest a similar distribution of 'Age at first kill' and 'Career-duration' for killers with "Escape" and "Angel of Death" motives of murder. But both seem to differ significantly with "Gang/Cult" killers.

Therefore, we require a hypothesis test to have some surety on whether age at first kill actually changes with motive.

Distribution – sample of "Age at first kill" of serial killers with different motives

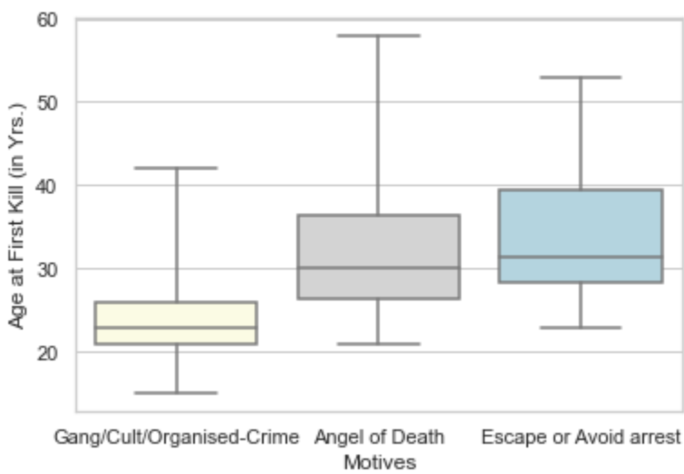


Figure 2 Box-plot of killer's "Age at first kill" for different motives

Figure 2 shows the boxplot of for all the three motives to get an idea about their distribution in comparison.

1. Distribution of killer’s data who are involved in Gang, cult or organised crime

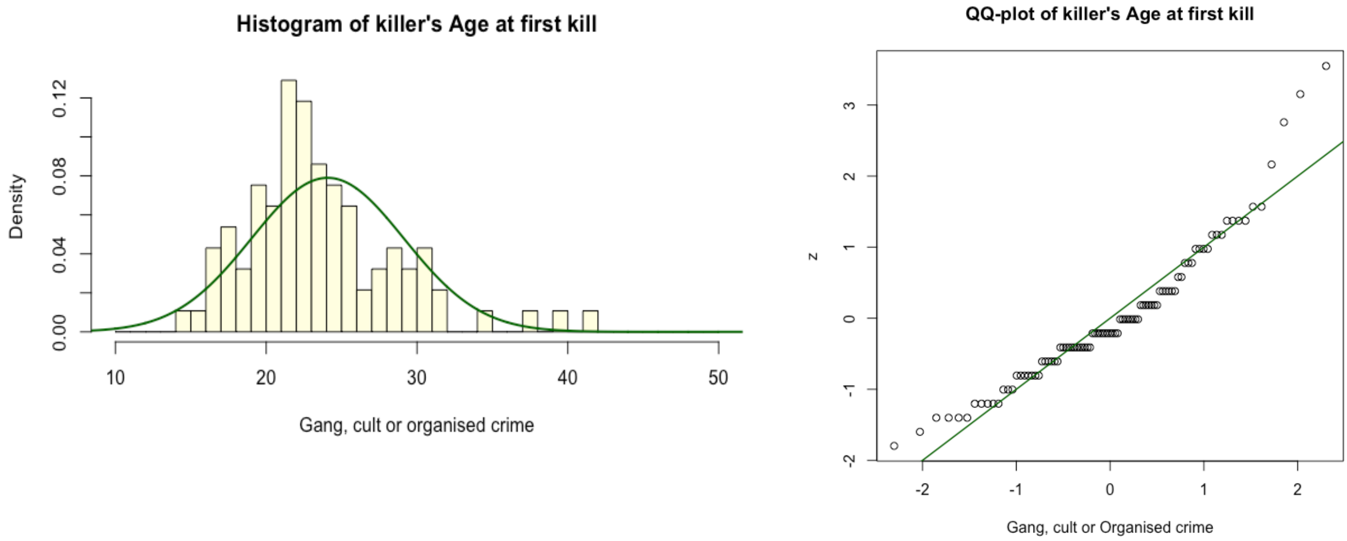


Figure 3. Histogram of killer’s Age at first kill (left), QQ-plot for normality (Right). (Motive ~ Gang, cult or Organised crime)

2. Distribution of killer’s data who are Angels of Death killers

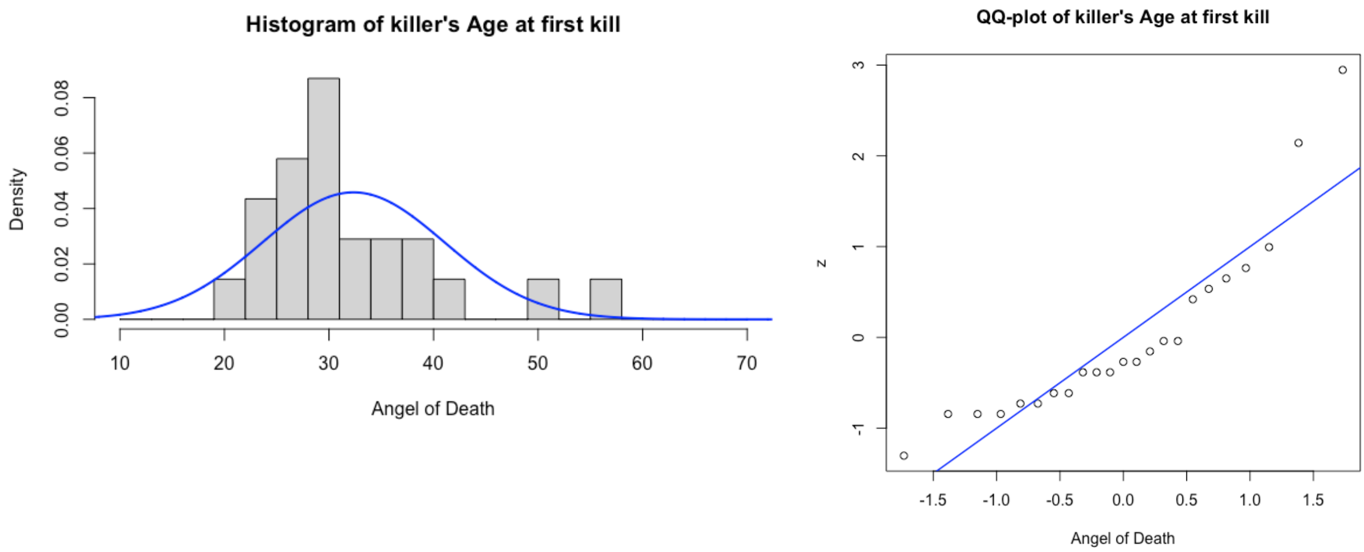


Figure 4. Histogram of killer’s Age at first kill (left), QQ-plot for normality (Right). (Motive ~ Angel of Death)

3. Distribution of killer’s data who kills to Escape or avoid arrest

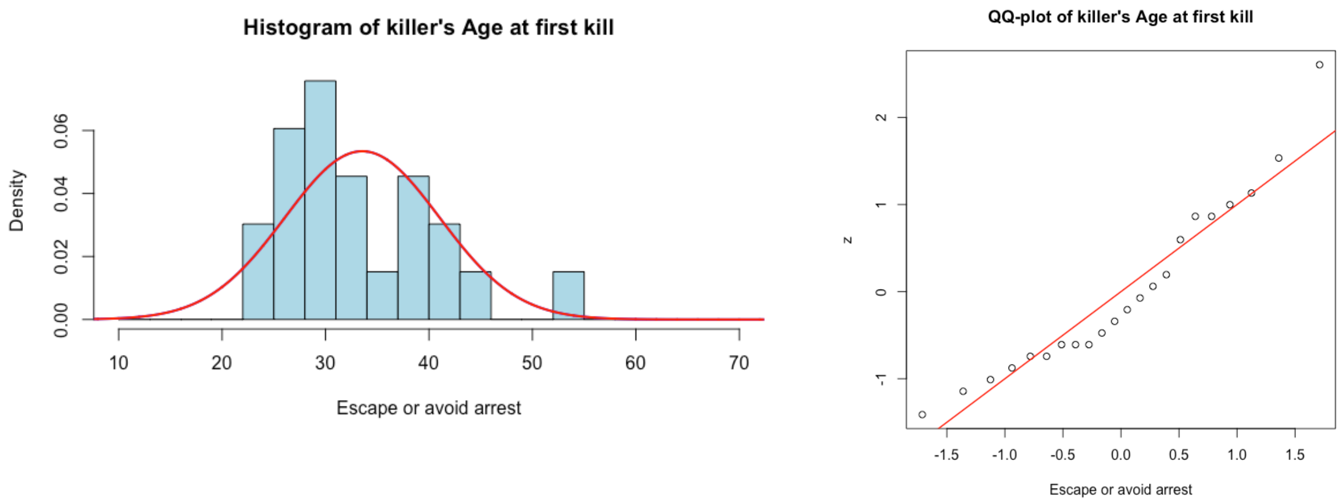


Figure 5. Histogram of killer’s Age at first kill (left), QQ-plot for normality (Right). (Motive ~ Escape or avoid arrest)

It can be seen from *Figure 3,4 and 5*, The Histogram shows the distribution fitting a normal curve and QQ-plot gives a supporting evidence of Normality for all the three samples. Therefore, Normality can be assumed for all of them.

Hypothesis tests I

First we consider Hypothesis tests to compare the population mean of killer's Age for different motives with that found by a previous study, which suggested population mean being equal to 27.
We do this for each sample of "Age at first kill by Motive".
Hypothesis taken for each test:

$$H_0: \mu = 27 \quad \textit{(Null Hypothesis)}$$
$$H_1: \mu \neq 27 \ \textit{(Alternative Hypothesis)}$$

where $'\mu'$ represents the population mean.

Tests needs to be performed on "Age at first kill" data for the following motives
Test 1. "Gang, cult, or Organised crime"
Test 2. "Angel of Death"
Test 3. "Escape or avoid arrest"

Assumptions taken for each test:
- The sample is considered as an independent and identically distributed (IID) sample.
- Population variance is not known
- Samples are assumed to come from a Normal distribution, see *Figure 3, 4 and 5.*

Based on the assumptions a t-test was performed, see *Table 2* below.

Test	Test sample	Sample Mean	Sig. level	Confidence-interval	p-value
1.	Gang, cult or Organised crime	24.07527	5%	23.03, 25.11	2.355e-07
2.	Angel of Death	32.34783	5%	28.58, 36.11	0.007
3.	Killed to escape or avoid arrest	33.54545	5%	30.23, 36.85	0.0005

Table 2. Results of Hypothesis tests performed for each sample of "Age at first kill" by Motive.

For each of the sample, we can see that there is a proof against the Null hypothesis. This suggests that the population mean found by previous study could be wrong.

Hypothesis tests II

We now intend to find out whether the Motive affects the Age at first kill for killers. Whether data for different motives have a different average Age of first kill.

Assumption of same population variance between each sample is made based on *Table 3* where 'A', 'B', 'C' represents the sample of "Age at first kill" for each motive.

	Sample STD.	1/3 < A/B < 3	1/3 < C/B < 3	1/3 < A/C < 3
Angel of Death (A)	8.705	True	-	-
Gang, cult .. (B)	5.05	-	True	-
Escape or avoid arrest (C)	7.46	-	-	True

Table 3. The test shows that we can assume same population variance between each sample.

Note,
μ_a - population mean of Age at first kill of killers with motive a.
μ_b - population mean of Age at first kill of killers with motive b.

Hypothesis taken for each test:

$$H_0: \mu_a - \mu_b = 0 \ \ (Null\ Hypothesis)$$
$$H_1: \mu_a - \mu_b \neq 0 \ \ (Alternative\ Hypothesis)$$

Test needs to be performed using “Age at first kill” data for motive ‘a’ and ‘b’, for the following combinations:

Test 1. a=“Angel of Death” & b=“Gang, cult, or Organised crime”
Test 2. a=“Angel of Death” & b=“Escape or avoid arrest”
Test 3. a=“Escape or avoid arrest” & b=“Gang, cult, or Organised crime”

Assumptions taken for both samples in the tests are:

- Both samples are considered as independent and identically distributed (IID) samples.
- Both samples are considered independent of each other.
- Population variance is not known but is considered same for both. See *Table 3*.
- Both samples are considered to have a Normal distribution, see *Figure 3, 4 and 5.*

Based on the assumptions t-test was performed for all the tests, see *Table 4* below.
In *Table 4,* Test-samples as

A – Angel of Death
B – Gang, cult or Organised crime
C – Escape or avoid arrest

Test	Test samples	Sample mean diff.	Sig. level	Confidence-Interval	p-value
1.	A & B	8.272	5%	5.535, 11.009	2.538e-08
2.	A & C	-1.197	5%	-6.084, 3.688	0.623
3.	C & B	9.470	5%	6.849, 12.090	8.747e-11

Table 4. Results of Hypothesis tests performed between samples of “Age at first kill” by Motive.

For test 1 and test 3 the expected difference of population mean (0) does not lie in the confidence interval and therefore we have proof against the null. Null can be rejected.

The only test where we have a proof supporting null hypothesis is test 2. The difference of population mean lies in the confidence interval. Therefore this null can be accepted.

Discussions

As for our main findings, then this can be summarised in the following points:

- Our analysis is in contradiction with the previous study which suggested that population mean of Age of killers at first kill was 27. The average age of killers at first kill seems to differ. It is lower than 27 for those who are involved in gangs and higher for those who kill to escape and those who kill people under their care. In other words, it suggests that serial killers involved in gangs start their career before the age of 27. And that serial killers who kill to escape or are Angels of death start much later around 28 years of age.
- Our analysis suggested that motives behind murder can lead to some killer start earlier than others.
- Another point suggested by this study is that those who kill people under their care (Angels of Death) and those who kill in order to escape, typically start their career (or kills their first target) around the same age. As for those who are involved in gangs then the study suggests that they start earlier when compared.

It must be noted that even though we have some evidences to support the assumptions of our hypothesis tests, some uncertainty still remains. This is Because the assumptions made about the actual data may or may not be true. In future with access to more data a much more certain result could be sought.

Appendix
R-Script for the analysis

This section contains the R-script which was written in order to achieve the Results mentioned above.

Part 1. Loading, Cleaning and Understanding the data.

```
load("killersandmotives.Rdata")
createsample(x=22)

## DATA CLEANING ##
# Clearing AgeFirstKill from special vals
tot_obs=length(mysample$AgeFirstKill)
mysample0=mysample[mysample$AgeFirstKill<150,]
summary(mysample0)

# Calculating Career Duration
mysample0$CareerDuration=mysample0$AgeLastKill - mysample0$AgeFirstKill

# Calculating Year at first kill - MySample
mysample$YearAtFirstKill=mysample$YearBorn + mysample$AgeFirstKill

# Calculating Year at first kill - MySample0
mysample0$YearAtFirstKill=mysample0$YearBorn + mysample0$AgeFirstKill
mysample0=mysample0[mysample0$YearAtFirstKill>1900,]
sum(mysample0$YearAtFirstKill<1900)

# Checking Columns for NA
for (i in names(mysample0)){
  if(sum(is.na(mysample0[i]))>0){
    print(i)
  }
}

## Motive, Sentence and InsanityPlea has NAs.
## Not considering Sentence and InsanityPlea as for now

# Removing NAs from Motive
sum(is.na(mysample0$Motive))
mysample0=mysample0[!is.na(mysample0$Motive),]
sum(is.na(mysample0$Motive))

# Getting observations removed
rem_obs=length(mysample0$AgeFirstKill)
del_obs=tot_obs - rem_obs
# Number of observations removed
print(del_obs)
# Percentage observations removed
(del_obs/tot_obs)*100

## DATA EXPLORATION ##
hist(mysample0$AgeFirstKill,freq=FALSE)
hist(mysample0$AgeLastKill,freq=FALSE)
hist(mysample0[mysample0$CareerDuration>0,"CareerDuration"])
summary(mysample0$AgeFirstKill)
```

```
summary(mysample0$AgeLastKill)
summary(mysample0$CareerDuration)
quantile(mysample0$AgeFirstKill,type=1)
quantile(mysample0$AgeLastKill,type=1)
quantile(mysample0[mysample0$CareerDuration>0,"CareerDuration"],type=1)
# Correlation
cor(mysample0$AgeFirstKill,mysample0$AgeLastKill)
plot(mysample0$AgeFirstKill,mysample0$AgeLastKill)

mysample1=mysample0[mysample0$CareerDuration>0,]
cor(mysample1$AgeFirstKill,mysample1$CareerDuration)
plot(mysample1$AgeFirstKill,mysample1$CareerDuration)

cor(mysample1$AgeLastKill,mysample1$CareerDuration)
plot(mysample1$AgeLastKill,mysample1$CareerDuration)
# Some additional Exploration // Effect of Motive
unique(mysample0$Motive)
gang=mysample0[mysample0$Motive=="Gang, cult or organised
crime",c("Motive","AgeFirstKill","AgeLastKill","CareerDuration")]
aod=mysample0[mysample0$Motive=="Angel of
Death",c("Motive","AgeFirstKill","AgeLastKill","CareerDuration")]
esc=mysample0[mysample0$Motive=="Escape or avoid
arrest",c("Motive","AgeFirstKill","AgeLastKill","CareerDuration")]

asian=mysample0[mysample0$Race=="Asian",c("Race","AgeFirstKill","AgeLastKill","CareerDuration")]
hisp=mysample0[mysample0$Race=="Hispanic",c("Race","AgeFirstKill","AgeLastKill","CareerDuration")
]
black=mysample0[mysample0$Race=="Black",c("Race","AgeFirstKill","AgeLastKill","CareerDuration")]
white=mysample0[mysample0$Race=="White",c("Race","AgeFirstKill","AgeLastKill","CareerDuration")]

male=mysample0[mysample0$Sex=="Male",c("Sex","AgeFirstKill","AgeLastKill","CareerDuration")]
female=mysample0[mysample0$Sex=="Female",c("Sex","AgeFirstKill","AgeLastKill","CareerDuration")]

length(aod$AgeFirstKill)
length(gang$AgeFirstKill)
length(esc$AgeFirstKill)

length(asian$AgeFirstKill)
length(hisp$AgeFirstKill)
length(black$AgeFirstKill)
length(white$AgeFirstKill)

length(male$AgeFirstKill)
length(female$AgeFirstKill)

print("### Hispanic ###")
print("Age First Kill")
quantile(hisp$AgeFirstKill,type=1)
IQR(hisp$AgeFirstKill,type=1)
print("Age Last Kill")
quantile(hisp$AgeLastKill,type=1)
IQR(hisp$AgeLastKill,type=1)
print("Career Duration")
quantile(hisp$CareerDuration,type=1)
IQR(hisp$CareerDuration,type=1)
```

```
print("### Asian ###")
print("Age First Kill")
quantile(asian$AgeFirstKill,type=1)
IQR(asian$AgeFirstKill,type=1)
print("Age Last Kill")
quantile(asian$AgeLastKill,type=1)
IQR(asian$AgeLastKill,type=1)
print("Career Duration")
quantile(asian$CareerDuration,type=1)
IQR(asian$CareerDuration,type=1)


print("### Black ###")
print("Age First Kill")
quantile(black$AgeFirstKill,type=1)
IQR(black$AgeFirstKill,type = 1)
print("Age Last Kill")
quantile(black$AgeLastKill,type=1)
IQR(black$AgeLastKill,type=1)
print("Career Duration")
quantile(black$CareerDuration,type=1)
IQR(black$CareerDuration,type=1)



print("### White ###")
print("Age First Kill")
quantile(white$AgeFirstKill,type=1)
IQR(white$AgeFirstKill,type = 1)
print("Age Last Kill")
quantile(white$AgeLastKill,type=1)
IQR(white$AgeLastKill,type=1)
print("Career Duration")
quantile(white$CareerDuration,type=1)
IQR(white$CareerDuration,type=1)


print("### Male ###")
print("Age First Kill")
quantile(male$AgeFirstKill,type=1)
IQR(male$AgeFirstKill,type = 1)
print("Age Last Kill")
quantile(male$AgeLastKill,type=1)
IQR(male$AgeLastKill,type=1)
print("Career Duration")
quantile(male$CareerDuration,type=1)
IQR(male$CareerDuration,type=1)


print("### Female ###")
print("Age First Kill")
quantile(female$AgeFirstKill,type=1)
```

```
IQR(female$AgeFirstKill,type = 1)
print("Age Last Kill")
quantile(female$AgeLastKill,type=1)
IQR(female$AgeLastKill,type=1)
print("Career Duration")
quantile(female$CareerDuration,type=1)
IQR(female$CareerDuration,type=1)

print("### AOD ###")
print("Age First Kill")
quantile(aod$AgeFirstKill,type=1)
IQR(aod$AgeFirstKill,type = 1)
print("Age Last Kill")
quantile(aod$AgeLastKill,type=1)
IQR(aod$AgeLastKill,type=1)
print("Career Duration")
quantile(aod$CareerDuration,type=1)
IQR(aod$CareerDuration,type=1)

print("### GANG ###")
print("Age First Kill")
quantile(gang$AgeFirstKill,type=1)
IQR(gang$AgeFirstKill,type = 1)
print("Age Last Kill")
quantile(gang$AgeLastKill,type=1)
IQR(gang$AgeLastKill,type=1)
print("Career Duration")
quantile(gang$CareerDuration,type=1)
IQR(gang$CareerDuration,type=1)

print("### ESC ###")
print("Age First Kill")
quantile(esc$AgeFirstKill,type=1)
IQR(esc$AgeFirstKill,type = 1)
print("Age Last Kill")
quantile(esc$AgeLastKill,type=1)
IQR(esc$AgeLastKill,type=1)
print("Career Duration")
quantile(esc$CareerDuration,type=1)
IQR(esc$CareerDuration,type=1)
```

Part 2. Investigating underlying distributions of 'Age at first kill' data for different Motives.

```
# AgeFirstKill – GANG, CULT OR ORGANISED CRIME
sample=gang
hist(sample$AgeFirstKill,freq=FALSE,breaks=seq(10,50,by=1),xlab=sample$Motive[1],main=paste("Histogram of killer's Age at first kill"),col="lightyellow",border=TRUE)
lines(seq(0,100,0.02),dnorm(seq(0,100,0.02),mean=mean(sample$AgeFirstKill),sd=sd(sample$AgeFirstKill)),lwd=2,col="darkgreen")
# KS Test ## PASSED
ks.test(x =  sample$AgeFirstKill, y = "pnorm", mean=mean(sample$AgeFirstKill), sd=sd(sample$AgeFirstKill))
# QQ-Plot test ## PASSED
mu <- mean(sample$AgeFirstKill)
```

```
sigma <- sd(sample$AgeFirstKill)
# Standardised order statistics:
z <- (sort(sample$AgeFirstKill) - mu)/sigma
n <- length(sample$AgeFirstKill)
r <- (1:n)
# Quantiles of N(0, 1):
q <- qnorm(p = r/(n + 1), mean = 0, sd = 1)
# A normal "QQ plot":
plot(q, z,xlab="Gang, cult or Organised crime", main="QQ-plot of killer's Age at first kill")
abline(a=0,b=1,col='darkgreen',lwd=1.5)
# Chi-squared test ## PASSED
library(nortest)
pearson.test(sample$AgeFirstKill)
## gang -> Normally distributed


# AgeFirstKill – ANGEL OF DEATH
sample=aod
hist(sample$AgeFirstKill,freq=FALSE,breaks=seq(10,70,by=3),xlab=sample$Motive[1],main=paste("Histo
gram of killer's Age at first kill"),col="lightgray",border=TRUE)
lines(seq(0,100,by=0.02),dnorm(seq(0,100,by=0.02),mean=mean(sample$AgeFirstKill),sd=sd(sample$Age
FirstKill)),lwd=2,col="blue")
# KS Test ## PASSED
ks.test(x =  sample$AgeFirstKill, y = "pnorm", mean=mean(sample$AgeFirstKill),
sd=sd(sample$AgeFirstKill))
# QQ-Plot test ## PASSED
mu <- mean(sample$AgeFirstKill)
sigma <- sd(sample$AgeFirstKill)
# Standardised order statistics:
z <- (sort(sample$AgeFirstKill) - mu)/sigma
n <- length(sample$AgeFirstKill)
r <- (1:n)
# Quantiles of N(0, 1):
q <- qnorm(p = r/(n + 1), mean = 0, sd = 1)
# A normal "QQ plot":
plot(q, z,xlab="Angel of Death", main="QQ-plot of killer's Age at first kill")
abline(a=0,b=1,col='blue',lwd=1.5)
# Chi-squared test ## PASSED
library(nortest)
pearson.test(sample$AgeFirstKill)
## aod - Normally Distributed


# AgeFirstKill -- ESCAPE OR AVOID ARREST
sample=esc
hist(sample$AgeFirstKill,freq=FALSE,breaks=seq(10,70,by=3),xlab=sample$Motive[1],main=paste("Histo
gram of killer's Age at first kill"),col="lightblue",border=TRUE)
lines(seq(0,100,by=0.02),dnorm(seq(0,100,by=0.02),mean=mean(sample$AgeFirstKill),sd=sd(sample$Age
FirstKill)),lwd=2,col="red")
lines(seq(0,100,by=0.02),dexp(seq(0,100,by=0.02),rate=1/mean(sample$AgeFirstKill)),lwd=2,col="red")
# KS Test ## PASSED
ks.test(x =  sample$AgeFirstKill, y = "pnorm", mean=mean(sample$AgeFirstKill),
sd=sd(sample$AgeFirstKill))
# QQ-Plot test ## PASSED
mu <- mean(sample$AgeFirstKill)
```

```
sigma <- sd(sample$AgeFirstKill)
# Standardised order statistics:
z <- (sort(sample$AgeFirstKill) - mu)/sigma
n <- length(sample$AgeFirstKill)
r <- (1:n)
# Quantiles of N(0, 1):
q <- qnorm(p = r/(n + 1), mean = 0, sd = 1)
# A normal "QQ plot":
plot(q, z,xlab="Escape or avoid arrest", main="QQ-plot of killer's Age at first kill")
abline(a=0,b=1,col='red',lwd=1.5)

# Chi-squared test ## PASSED
library(nortest)
pearson.test(sample$AgeFirstKill)

# esc - Normally distributed

## Up until-now it has been seen that all the three datasets
## are Normally distributed. therefore it is assumed that they
## come from normally distributed population.
```

Part 3. Quality check & Hypothesis tests I.

```
# Quality check
for (i in c('Angel of Death','Gang, cult or organised crime','Escape or avoid arrest')){
  len0=length(mysample[mysample$Motive==i,"AgeFirstKill"])
  nulls0=sum(is.na(mysample[mysample$Motive==i,"AgeFirstKill"]))
  specials0=length(mysample[mysample$Motive==i & mysample$AgeFirstKill>110,"AgeFirstKill"])
  years0=length(mysample[mysample$Motive==i & mysample$YearAtFirstKill<1900,"AgeFirstKill"])
  expectedlength=len0-nulls0-specials0-years0
  if (i=="Angel of Death"){
    sample=aod
  }
  else if(i=="Gang, cult or organised crime"){
    sample=gang
  }
  else{
    sample=esc
  }
  flag=length(sample$AgeFirstKill)==expectedlength
  if (flag){
    print(c("All seems good for",i))
  }
  else{
    print(c("Issues with",i))
  }
}


boxplot(mysample0$AgeFirstKill ~ mysample0$Motive,type=1, range=5) # Range to extend the box


# Hypothesis testing
# Is pop. mean of aod, esc and gang equal to 27?
# defining a function to reduce hassle
ztest=function(x,mu,sigma2=sd(x)^2,siglevel){
```

```
  xbar=mean(x)
  n=length(x)
  alphaby2=siglevel/2
  CI=xbar + c(-1,1)*qnorm(p=(1-alphaby2),mean=0,sd=1)*sqrt(sigma2/n)
  z=(xbar-mu)/sqrt(sigma2/n)
  pval=2*pnorm(q=-z,mean=0,sd=1)
  return(c(CI,pval,z))
}

ttest=function(x,mu,siglevel){
  xbar=mean(x)
  n=length(x)
  samplevar=sd(x)^2
  alphaby2=siglevel/2
  CI=xbar + c(-1,1)*qt(p=(1-alphaby2),df=n-1)*sqrt(samplevar/n)
  T=(xbar-mu)/sqrt(samplevar/n)
  pval=2*pt(q=-(T),df=n-1)
  print(c("CI",CI))
  print(c("t",T))
  print(c("p-value",pval))
}

# ---- # # ---- # # ---- # # ---- # # ---- # # ---- # # ---- #
## MAIN ## For Angel of Death (aod)
# Ho: mu=27 Vs H1: mu!=27

# Assumptions
# 1. the sample is assumed as an IID sample
# 2. Normal distribution is assumed (supported by our test above)
# 3. sigma^2 unknown (sigma2=samplevar)
# Therefore we proceed with t-test

x=aod$AgeFirstKill
ttest(x,mu=27,siglevel=0.05)
# CI: (28.58,36.11) does'nt contain 27
t.test(x,mu=27,siglevel=0.05) # Confirmation

# Therefore proof against null

# ---- # # ---- # # ---- # # ---- # # ---- # # ---- # # ---- #
## MAIN ## For Gang, cult or organised crime (gang)
# Ho: mu=27 Vs H1: mu!=27

# (this is actual) Assumptions
# 1. the sample is assumed as an IID sample
# 2. Normal distribution is assumed (supported by our test above)
# 3. sigma^2 unknown (sigma2=samplevar)
# Therefore we proceed with t-test

x=gang$AgeFirstKill
ttest(x,mu=27,siglevel=0.05)
# CI: (23.035,25.115) does'nt contain 27
t.test(x,mu=27) # Confirmation

# Therefore proof against null
```

```
# ---- # # ---- # # ---- # # ---- # # ---- # # ---- # # ---- #
## MAIN ## For Escape or avoid arrest
# Ho: mu=27 Vs H1: mu!=27

# (this is actual) Assumptions
# 1. the sample is assumed as an IID sample
# 2. Normal distribution is assumed (supported by our test above)
# 3. sigma^2 unknown (sigma2=samplevar)
# Therefore we proceed with t-test

x=esc$AgeFirstKill
ttest(x,mu=27,siglevel=0.05)
# CI: (30.23,36.85) does'nt contain 27
t.test(x,mu=27) # Confirmation
# Therefore proof against null


# Therefore proof against null

# ---- # # ---- # # ---- # # ---- # # ---- # # ---- # # ---- #
# ---- # # ---- # # ---- # # ---- # # ---- # # ---- # # ---- #
```

Part 4. Hypothesis tests II.

```
# Ho: mu(aod)-mu(gang)=0 Vs H1: mu(aod)-mu(gang)!=0
x1=aod$AgeFirstKill
x2=gang$AgeFirstKill

# Assumptions
  # 1. Normality is assumed for both the samples
  # 2. both are independent samples and are independent within themselves
1/3<sd(x1)/sd(x2)
sd(x1)/sd(x2)<3
  # 3. based on the check above The pop. variance are considered same for both
  # 4. Significance level is 0.05

x1barx2bar=mean(x1)-mean(x2)
s12=sd(x1)^2
s22=sd(x2)^2
n1=length(x1)
n2=length(x2)
serror=sqrt((1/n1)+(1/n2))
df=n1+n2-2
tn1n2=qt(0.975,df=df)
sp2=((n1-1)*s12 + (n2-1)*s22)/(n1+n2-2)

T=x1barx2bar/(sqrt(sp2)*serror)
CI=x1barx2bar+c(-1,1)*tn1n2*sqrt(sp2)*serror
pval=2*pt(q=-T,df=df)
t.test(x=x1,y=x2,mu=0,var.equal=TRUE,paired=FALSE) # Confirmation
# The avg. age seem to differ between killers with motive
# Angel of Death and those associated with Gang, cult or organised crime

# --- --- # --- --- # --- --- # # --- --- # --- --- # --- --- #
```

```
# Ho: mu(aod)-mu(esc)=0 Vs H1: mu(aod)-mu(esc)!=0
x1=aod$AgeFirstKill
x2=esc$AgeFirstKill

# Assumptions
# 1. Normality is assumed for both the samples
# 2. both are independent samples and are independent within themselves
1/3<sd(x1)/sd(x2)
sd(x1)/sd(x2)<3
# 3. based on the check above The pop. variance are considered same for both
# 4. Significance level is 0.05

x1barx2bar=mean(x1)-mean(x2)
s12=sd(x1)^2
s22=sd(x2)^2
n1=length(x1)
n2=length(x2)
serror=sqrt((1/n1)+(1/n2))
df=n1+n2-2
tn1n2=qt(0.975,df=df)
sp2=((n1-1)*s12 + (n2-1)*s22)/(n1+n2-2)

T=x1barx2bar/(sqrt(sp2)*serror)
pval=2*pt(q=-T,df=df)
CI=x1barx2bar+c(-1,1)*tn1n2*sqrt(sp2)*serror
t.test(x=x1,y=x2,mu=0,var.equal=TRUE) # Confirmation
# The avg. age does'nt seem to differ between killers with motive
# Angel of Death and those associated with Escape or avoid arrest

# --- --- # --- --- # --- --- # # --- --- # --- --- # --- --- #

# Ho: mu(aod)-mu(esc)=0 Vs H1: mu(aod)-mu(esc)!=0
x1=esc$AgeFirstKill
x2=gang$AgeFirstKill

# Assumptions
# 1. Normality is assumed for both the samples
# 2. both are independent samples and are independent within themselves
1/3<sd(x2)/sd(x1)
sd(x2)/sd(x1)<3
# 3. based on the check above The pop. variance are considered same for both
# 4. Significance level is 0.05

x1barx2bar=mean(x1)-mean(x2)
s12=sd(x1)^2
s22=sd(x2)^2
n1=length(x1)
n2=length(x2)
serror=sqrt((1/n1)+(1/n2))
df=n1+n2-2
tn1n2=qt(0.975,df=df)
sp2=((n1-1)*s12 + (n2-1)*s22)/(n1+n2-2)

T=x1barx2bar/(sqrt(sp2)*serror)
```

```
pval=2*pt(q=-T,df=df)
CI=x1barx2bar+c(-1,1)*tn1n2*sqrt(sp2)*serror
t.test(x=x1,y=x2,mu=0,var.equal=TRUE) # Confirmation
# The avg. age seem to differ between killers with motive
# Escape or avoid arrest and those associated with Gang, cult or organised crime
```