**Please do not change the coding scheme, otherwise will not give correct values as it is documented in the report.**

**Libraries**

|  |
| --- |
| library(tidyverse) library(dplyr) library(ggplot2) library(corrplot) library(lubridate) |

**Distribution graphs**

|  |
| --- |
| hist(london\_merged$cnt) hist(london\_merged$t1) hist(london\_merged$t2) hist(london\_merged$hum) hist(london\_merged$wind\_speed) hist(london\_merged$weather\_code) hist(london\_merged$is\_weekend) hist(london\_merged$is\_holiday) hist(london\_merged$season) |

**Boxplot to see the outliers**

|  |
| --- |
| boxplot(london\_merged$cnt, col="slateblue") x\_out\_rm <- london\_merged$cnt[!london\_merged$cnt %in% boxplot.stats(london\_merged$cnt)$out] |

**Boxplot**

|  |
| --- |
| length(london\_merged$cnt) - length(x\_out\_rm) boxplot(x\_out\_rm) boxplot(x\_out\_rm, horizontal = TRUE, axes = FALSE, staplewex = 1) text(x=fivenum(x\_out\_rm), labels =fivenum(x\_out\_rm), y=1.25) |

**Removing outliers**

|  |
| --- |
| range\_cnt <- quantile(london\_merged\_test$cnt, probs = c(0.25, 0.75), na.rm = FALSE) interQuartile\_cnt <- IQR(london\_merged\_test$cnt)  Q\_hum <- quantile(london\_merged\_test$hum, probs = c(0.25, 0.75), na.rm = FALSE) iqr\_hum <- IQR(london\_merged\_test$hum)  range\_wind\_speed <- quantile(london\_merged\_test$wind\_speed, probs = c(0.25, 0.75), na.rm = FALSE) interQuartile\_wind\_speed <- IQR(london\_merged\_test$wind\_speed)  high\_cnt <- range\_cnt[2]+1.5\*interQuartile\_cnt low\_cnt <- range\_cnt[1]-1.5\*interQuartile\_cnt  up\_hum <- Q\_hum[2]+1.5\*iqr\_hum low\_hum <- Q\_hum[1]-1.5\*iqr\_hum  high\_wind\_speed <- range\_wind\_speed [2]+1.5\*interQuartile\_wind\_speed low\_wind\_speed <- range\_wind\_speed [1]-1.5\*interQuartile\_wind\_speed  london\_merged\_test=subset(london\_merged\_test, cnt > low\_cnt & cnt < high\_cnt & wind\_speed > low\_wind\_speed & wind\_speed < high\_wind\_speed) |

**Correlation Matrix**

|  |
| --- |
| library(corrplot) library("dplyr") NUMDATA2<-select\_if(london\_merged\_test, is.numeric) corrplot(cor(NUMDATA2), type="lower", method="number") |

**Linear regression graphs**

|  |
| --- |
| # `cnt` vs `t1`  x <- london\_merged\_test$t1 y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Actual Temperature",  xlab = "Actual Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Actual Temperature",  xlab = "Actual Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `t2`  x <- london\_merged\_test$t2 y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Feels Like Temperature",  xlab = "Feels Like Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Feels Like Temperature",  xlab = "Feels Like Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `hum`  x <- london\_merged\_test$hum y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Humidity",  xlab = "Humidity", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Humidity",  xlab = "Humidity", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `weather\_code`  x <- london\_merged\_test$weather\_code y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Weather Code",  xlab = "Weather Code", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Weather Code",  xlab = "Weather Code", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `is\_holiday`  x <- london\_merged\_test$is\_holiday y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Is Holiday",  xlab = "Is Holiday", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Is Holiday",  xlab = "Is Holiday", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `is\_weekend`  x <- london\_merged\_test$is\_weekend y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Is Weekend",  xlab = "Is Weekend", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Is Weekend",  xlab = "Is Weekend", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `season`  x <- london\_merged\_test$season y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Season",  xlab = "Season", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Season",  xlab = "Season", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

**Regression graphs for independent vs dependent variables**

|  |
| --- |
| # `cnt` vs `t1`  x <- london\_merged\_test$t1 y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Actual Temperature",  xlab = "Actual Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Actual Temperature",  xlab = "Actual Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `t2`  x <- london\_merged\_test$t2 y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Feels Like Temperature",  xlab = "Feels Like Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Feels Like Temperature",  xlab = "Feels Like Temperature", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `hum`  x <- london\_merged\_test$hum y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Humidity",  xlab = "Humidity", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Humidity",  xlab = "Humidity", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `wind\_speed`  x <- london\_merged\_test$wind\_speed y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Wind Speed",  xlab = "Wind Speed", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Wind Speed",  xlab = "Wind Speed", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `weather\_code`  x <- london\_merged\_test$weather\_code y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Weather Code",  xlab = "Weather Code", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Weather Code",  xlab = "Weather Code", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `is\_holiday`  x <- london\_merged\_test$is\_holiday y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Is Holiday",  xlab = "Is Holiday", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Is Holiday",  xlab = "Is Holiday", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `is\_weekend`  x <- london\_merged\_test$is\_weekend y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Is Weekend",  xlab = "Is Weekend", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Is Weekend",  xlab = "Is Weekend", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

|  |
| --- |
| # `cnt` vs `season`  x <- london\_merged\_test$season y <- london\_merged\_test$cnt  plot(x, y, main = "New Bike Share VS Season",  xlab = "Season", ylab = "Bike Share Counts",  pch = 19, frame = FALSE)  plot(x, y, main = "New Bike Share VS Season",  xlab = "Season", ylab = "Bike Share Counts",  pch = 19, frame = FALSE) abline(lm(y ~ x, data = london\_merged\_test), col = "blue") |

**Frequency graphs**

**`cnt` vs `is\_holiday`data <- london\_merged\_test**

|  |
| --- |
| library(wesanderson) library(ggplot2) data$is\_holiday<-factor(data$is\_holiday, levels=c("0", "1"), labels=c("Non-Holiday", "Holiday"))  ggplot(data, aes(x=is\_holiday, y=cnt, fill=is\_holiday)) + geom\_boxplot() +  ggtitle("New bike counts Between is\_holiday") + xlab("Holiday") + ylab("New Bike Counts") +  theme(legend.position = "none") +  scale\_fill\_manual(values = wes\_palette("GrandBudapest1", type=c("discrete"), n=2))  t.test(data$cnt~data$is\_holiday) |

**`cnt` vs `is\_weekend`**

|  |
| --- |
| data$is\_weekend<-factor(data$is\_weekend, levels=c("0", "1"), labels=c("Non-Weekend", "Weekend"))  ggplot(data, aes(x=is\_weekend, y=cnt, fill=is\_weekend)) + geom\_boxplot() +  ggtitle("New bike counts Between is\_weekend") + xlab("is\_weekend") + ylab("New Bike Counts") +  theme(legend.position = "none") +  scale\_fill\_manual(values = wes\_palette("GrandBudapest1", type=c("discrete"), n=2))  t.test(data$cnt~data$is\_weekend) |

**`cnt` vs `season` boxplot**

|  |
| --- |
| library(magrittr) library(ggplot2) library(dplyr) dw =c("darkgreen", "darkgreen", "darkgreen", "darkgreen",  "darkgreen", "olivedrab3", "olivedrab3") b <- london\_merged\_test b <- b %>% rename("Rentals" = "cnt",  "Season" = 'season',  'Weather' = 'weather\_code',  'Workday' = 'is\_weekend')  b <- b %>% select("Rentals", "Temperature", "Feels like Temperature", "Humidity", "timestamp",  Season, Weather, Windspeed, Workday)  b$Season <- factor(  b$Season, levels = c(0,1,2,3),  labels = c('Spring', 'Summer', 'Fall','Winter'),  ordered = TRUE) (coefs <- coef(lm(cnt ~ season, data = london\_merged\_test))) ggplot(b,aes(Season, Rentals)) +  geom\_boxplot(aes(fill=Season), show.legend = F) +  theme\_bw(base\_size = 16) + scale\_fill\_manual(values = dw) +  labs(title = "New bike sharing counts by Season", x = "", y = "Count of new bike sharing counts") +  scale\_y\_continuous(labels = scales::label\_comma()) + geom\_abline(intercept = coefs[1], slope = coefs[2], color = "red") |

**`cnt` vs `weather\_code` boxplot**

|  |
| --- |
| b$Weather <- factor(  b$Weather,  levels = c(1,2,3,4,7,10,26),  labels = c('Clear', 'Scattered Clouds','Broken Clouds', 'Cloudy', 'Light Rain',  'Rain/Snow', 'Rain/Snow')) (coefs2 <- coef(lm(cnt ~ weather\_code, data = london\_merged\_test))) options(repr.plot.width=12, repr.plot.height=6) ggplot(b,aes(Weather, Rentals)) +  geom\_boxplot(aes(fill=Weather), show.legend = F) +  theme\_bw(base\_size = 16) + scale\_fill\_manual(values = dw) +  labs(title = "New bike sharing counts by Weather", x = "", y = "Count of new bike sharing counts") +  scale\_y\_continuous(labels = scales::label\_comma()) + geom\_abline(intercept = coefs[1], slope = coefs[2], color = "red") coefs2  boxplot(Rentals ~ Weather, data = b, frame = F, col = "grey40", main = "New bike sharing counts by weather\_code") |

**`cnt` vs `season` and `is\_weekend`**

|  |
| --- |
| library(magrittr) library(dplyr) b$Workday <- factor(b$Workday,  levels = c(0,1),  labels = c('Workday', 'Weekend')) s = c("olivedrab3", 'yellow', 'orange', 'grey50') options(repr.plot.width=12, repr.plot.height=8) options(repr.plot.width=12, repr.plot.height=8) b %>% group\_by(Season, Workday) %>%  summarise(n = n(), rent = sum(Rentals)) %>%  ggplot(aes(Season, rent, fill = Season)) +  geom\_bar(stat = "identity", color = 'grey50', show.legend = F) +  scale\_fill\_manual(values = s) +  facet\_grid(~Workday) + theme\_minimal(base\_size = 16) +  labs(title = "New bike sharing by season & is\_weekend", x = "", y = "New bike sharing counts") +  scale\_y\_continuous(labels = scales::label\_comma()) |

**Linear regression intercept and rate**

|  |
| --- |
| lm(london\_merged\_test$cnt ~ london\_merged\_test$t1) lm(london\_merged\_test$cnt ~ london\_merged\_test$t2) lm(london\_merged\_test$cnt ~ london\_merged\_test$hum) lm(london\_merged\_test$cnt ~ london\_merged\_test$wind\_speed) |

**Probability theory computation**

|  |
| --- |
| # Main #  n <- 50 r <- 5 m <- 2  a <- choose(n, r) b <- choose(r, m) c <- choose((n-r), (r-m))  f <- (b\*c)/a f  # Lucky Stars #  t <- 12 b <- 2 d <- 0  w <- choose(t, b) x <- choose(b, d) y <- choose((t-b), (b-d)) z <- (x\*y)/w z  z\*f |

**Splitting dataset into training and testing**

|  |
| --- |
| data2 <- london\_merged set.seed(1000) r <- runif(nrow(data2)) london\_reg <- data2[order(r), ] train <- london\_reg[1:13932, ] test <- london\_reg[13933:17414, ] str(train) str(test) |

**Training regression for `cnt` vs `t1`**

|  |
| --- |
| x <- train$t1 y <- train$cnt cor(train$cnt, train$t1) london\_training <- lm(cnt ~ t1, data = train) summary(london\_training) plot(train$t1, train$cnt, main ="cnt vs t1 - Training set", xlab = "Actual Temperatures", ylab = "New bike count", xlim=c(0, 35), ylim=c(0, 8000)) abline(lm(y ~ x, data = train), col = "blue") |

**Testing regression for `cnt` vs `t1`**

|  |
| --- |
| x <- test$t1 y <- test$cnt cor(test$cnt, test$t1) london\_training <- lm(cnt ~ t1, data = test) summary(london\_training) plot(test$t1, test$cnt, main ="cnt vs t1 - Testing set", xlab = "Actual Temperatures", ylab = "New bike count", xlim=c(0, 35), ylim=c(0, 8000)) abline(lm(y ~ x, data = test), col = "blue") |

**RSME required libraries**

|  |
| --- |
| library(Metrics) library(dplyr) |

**MAPE, RSME and MAE calculation for `cnt` vs `t1`**

|  |
| --- |
| london\_rmse <- train %>% mutate(predict.train = predict(london\_training)) london\_rmse\_test <- predict(london\_training, newdata = test) x <- london\_rmse$cnt y <- london\_rmse$predict.train plot(london\_rmse$cnt, london\_rmse$predict.train, main ="cnt vs t1 - Original vs Predicted", xlab = "Actual Counts", ylab = "Predicted Counts") abline(lm(y ~ x, data = london\_rmse), col = "blue") cor(london\_rmse$cnt, london\_rmse$predict.train)  mape\_train1 <- mape(london\_rmse$predict.train, train$cnt) mape\_train1 mape\_test1 <- mape(london\_rmse\_test, test$cnt) mape\_test1  rmse1 <- rmse(train$cnt, london\_rmse$predict.train) rmse rmse\_test <- rmse(test$cnt, london\_rmse\_test) rmse\_test  mae\_train1 <- mae(train$cnt, london\_rmse$predict.train) mae\_train1 mae\_test1 <- mae(test$cnt, london\_rmse\_test) mae\_test1 |

**Training regression for `cnt` vs `hum`**

|  |
| --- |
| x <- train$hum y <- train$cnt cor(train$cnt, train$hum) london\_training2 <- lm(cnt ~ hum, data = train) summary(london\_training2) #plot(london\_training2) plot(train$hum, train$cnt, main ="cnt vs hum - Training set", xlab = "Humidity", ylab = "New bike count") abline(lm(y ~ x, data = train), col = "blue") |

**Testing regression for `cnt` vs `hum`**

|  |
| --- |
| x <- test$hum y <- test$cnt cor(test$cnt, test$hum) london\_training <- lm(cnt ~ hum, data = test) summary(london\_training) plot(test$hum, test$cnt, main ="cnt vs hum - Testing set", xlab = "Actual Temperatures", ylab = "New bike count") abline(lm(y ~ x, data = test), col = "blue") |

**MAPE, RSME and MAE calculation for `cnt` vs `hum`**

|  |
| --- |
| london\_rmse <- train %>% mutate(predict.train2 = predict(london\_training2)) london\_rmse\_test2 <- predict(london\_training2, newdata = test) x <- london\_rmse$cnt y <- london\_rmse$predict.train2 plot(london\_rmse$cnt, london\_rmse$predict.train2, main ="cnt vs hum - Original vs Predicted", xlab = "Actual counts", ylab = "Predicted counts") abline(lm(y ~ x, data = london\_rmse), col = "blue") cor(london\_rmse$cnt, london\_rmse$predict.train2) summary(london\_rmse$predict.train2)  mape\_train2 <- mape(london\_rmse$predict.train2, train$cnt) mape\_train2 mape\_test2 <- mape(london\_rmse\_test2, test$cnt) mape\_test2  rmse2 <- rmse(train$cnt, london\_rmse$predict.train2) rmse2 rmse\_test2 <- rmse(test$cnt, london\_rmse\_test2) rmse\_test2  mae\_train2 <- mae(train$cnt, london\_rmse$predict.train2) mae\_train2 mae\_test2 <- mae(test$cnt, london\_rmse\_test2) mae\_test2 |

**Training regression and frequency graph for `cnt` vs `weather\_code`**

|  |
| --- |
| x <- train$weather\_code y <- train$cnt cor(train$cnt, train$weather\_code) london\_training3 <- lm(cnt ~ weather\_code, data = train) summary(london\_training3) plot(train$weather\_code, train$cnt, main ="cnt vs weather\_code - Training set", xlab = "Weather code", ylab = "New bike count") abline(lm(y ~ x, data = train), col = "blue")  train$weather\_code <- factor(  train$weather\_code,  levels = c(1,2,3,4,7,10,26),  labels = c('Clear', 'Scattered Clouds','Broken Clouds', 'Cloudy', 'Light Rain',  'Rain/Snow', 'Rain/Snow')) options(repr.plot.width=12, repr.plot.height=6) ggplot(train,aes(weather\_code, cnt)) +  geom\_boxplot(aes(fill=weather\_code), show.legend = F) +  theme\_bw(base\_size = 16) + scale\_fill\_manual(values = dw) +  labs(title = "New bike sharing counts by Weather", x = "", y = "Count of new bike sharing counts") +  scale\_y\_continuous(labels = scales::label\_comma()) |

**Testing regression and frequency graph for `cnt` vs `weather\_code`**

|  |
| --- |
| x <- test$weather\_code y <- test$cnt cor(test$cnt, test$weather\_code) london\_training <- lm(cnt ~ weather\_code, data = test) summary(london\_training) plot(test$weather\_code, test$cnt, main ="cnt vs weather\_code - Testing set", xlab = "Actual Temperatures", ylab = "New bike count") abline(lm(y ~ x, data = test), col = "blue")  test$weather\_code <- factor(  test$weather\_code,  levels = c(1,2,3,4,7,10,26),  labels = c('Clear', 'Scattered Clouds','Broken Clouds', 'Cloudy', 'Light Rain',  'Rain/Snow', 'Rain/Snow')) options(repr.plot.width=12, repr.plot.height=6) ggplot(test,aes(weather\_code, cnt)) +  geom\_boxplot(aes(fill=weather\_code), show.legend = F) +  theme\_bw(base\_size = 16) + scale\_fill\_manual(values = dw) +  labs(title = "New bike sharing counts by Weather", x = "", y = "Count of new bike sharing counts") +  scale\_y\_continuous(labels = scales::label\_comma()) |

**MAPE, RSME and MAE calculation for `cnt` vs `weather\_code`**

|  |
| --- |
| london\_rmse <- train %>% mutate(predict.train3 = predict(london\_training3)) london\_rmse\_test3 <- predict(london\_training3, newdata = test) x <- london\_rmse$cnt y <- london\_rmse$predict.train3 plot(london\_rmse$cnt, london\_rmse$predict.train3, main ="cnt vs weather\_code - Original vs Predicted", xlab = "Actual counts", ylab = "Predicted counts") abline(lm(y ~ x, data = london\_rmse), col = "blue") cor(london\_rmse$cnt, london\_rmse$predict.train3)  mape\_train3 <- mape(london\_rmse$predict.train3, train$cnt) mape\_train3 mape\_test3 <- mape(london\_rmse\_test3, test$cnt) mape\_test3  rmse3 <- rmse(train$cnt, london\_rmse$predict.train3) rmse3 rmse\_test3 <- rmse(test$cnt, london\_rmse\_test3) rmse\_test3  mae\_train3 <- mae(train$cnt, london\_rmse$predict.train3) mae\_train3 mae\_test3 <- mae(test$cnt, london\_rmse\_test3) mae\_test3 |

**Training regression for `cnt` with multiple variables**

|  |
| --- |
| london\_training4 <- lm(cnt ~ season+is\_holiday+t1+t2, data = train) summary(london\_training4) plot(london\_training4) res <- resid(london\_training4) plot(density(res)) |

**MAPE, RSME and MAE calculation for `cnt` vs multiple variables**

|  |
| --- |
| london\_rmse <- train %>% mutate(predict.train4 = predict(london\_training4)) london\_rmse\_test4 <- predict(london\_training4, newdata = test) x <- london\_rmse$cnt y <- london\_rmse$predict.train4 plot(london\_rmse$cnt, london\_rmse$predict.train4, main ="cnt vs multiple variables - Original vs Predicted", xlab = "Actual counts", ylab = "Predicted counts") abline(lm(y ~ x, data = london\_rmse), col = "blue") cor(london\_rmse$cnt, london\_rmse$predict.train4) mape\_train4 <- mape(london\_rmse$predict.train4, train$cnt) mape\_train4 mape\_test4 <- mape(london\_rmse\_test4, test$cnt) mape\_test4  rmse4 <- rmse(train$cnt, london\_rmse$predict.train4) rmse4 rmse\_test4 <- rmse(test$cnt, london\_rmse\_test4) rmse\_test4  mae\_train4 <- mae(train$cnt, london\_rmse$predict.train4) mae\_train4 mae\_test4 <- mae(test$cnt, london\_rmse\_test4) mae\_test4 |