black_scholes_final

Black Scholes

```
library(PerformanceAnalytics)
library(dplyr)
library(tidyverse)
```

Separating Call Options and Put Options

```
options_bs <- read.csv("msft_final_df2.csv")
head(options_bs)
```

```
##      X       date     exdate cp_flag strike_price best_bid best_offer volume
## 1 3867 2010-02-01 2012-01-21       C        37500     1.17       1.25    599
## 2 3868 2010-02-01 2010-02-20       P        29000     1.05       1.07   3739
## 3 3869 2010-02-01 2010-03-20       P        27000     0.51       0.53   5824
## 4 3870 2010-02-01 2012-01-21       C        22500     7.20       7.50     22
## 5 3871 2010-02-01 2010-03-20       P        25000     0.17       0.18    432
## 6 3872 2010-02-01 2010-03-20       P        26000     0.29       0.31    351
##   open_interest impl_volatility date_ndiff treasury_rate closing_price
## 1         41146        0.246200        719          0.86         28.41
## 2         29228        0.251661         19          0.05         28.41
## 3          4284        0.265746         47           0.1         28.41
## 4           595        0.288945        719          0.86         28.41
## 5          1410        0.301554         47           0.1         28.41
## 6          2364        0.281168         47           0.1         28.41
##     sigma_20
## 1 0.01521071
## 2 0.01521071
## 3 0.01521071
## 4 0.01521071
## 5 0.01521071
## 6 0.01521071
```

```
options_both <- options_bs[, c("date", "exdate", "cp_flag", "strike_price", "best_bid", "best_of
fer", "volume", "open_interest", "impl_volatility", "date_ndiff", "treasury_rate", "closing_pric
e", "sigma_20")]

# dropping columns impl_volatility and exdate
options_both[, 9] <- NULL
options_both[, 2] <- NULL

# dropping NA values from treasury_rate column
options_both$treasury_rate <- as.numeric(options_both$treasury_rate)
summary(options_both$treasury_rate)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0000  0.0700  0.4400  0.8718  1.6900  2.9800     167
```

```
nrow(options_both)
```

```
## [1] 1489016
```

```
options_both <- na.omit(options_both)
nrow(options_both)
```

```
## [1] 1488849
```

```
options_put <- options_both %>% filter(cp_flag == "P")
options_put[, 2] <- NULL
nrow(options_put)
```

```
## [1] 772771
```

```
head(options_put)
```

```
##         date strike_price best_bid best_offer volume open_interest date_ndiff
## 1 2010-02-01        29000     1.05       1.07   3739         29228         19
## 2 2010-02-01        27000     0.51       0.53   5824          4284         47
## 3 2010-02-01        25000     0.17       0.18    432          1410         47
## 4 2010-02-01        26000     0.29       0.31    351          2364         47
## 5 2010-02-01        28000     0.86       0.89   6838          2956         47
## 6 2010-02-01        26000     0.08       0.09     89          8513         19
##   treasury_rate closing_price   sigma_20
## 1          0.05         28.41 0.01521071
## 2          0.10         28.41 0.01521071
## 3          0.10         28.41 0.01521071
## 4          0.10         28.41 0.01521071
## 5          0.10         28.41 0.01521071
## 6          0.05         28.41 0.01521071
```

```
options_call <- options_both %>% filter(cp_flag == "C")
options_call[, 2] <- NULL
nrow(options_call)
```

```
## [1] 716078
```

```
head(options_call)
```

```
##         date strike_price best_bid best_offer volume open_interest date_ndiff
## 1 2010-02-01        37500     1.17       1.25    599         41146        719
## 2 2010-02-01        22500     7.20       7.50     22           595        719
## 3 2010-02-01        35000     1.62       1.76      0          2068        719
## 4 2010-02-01        45000     0.40       0.46      0          1697        719
## 5 2010-02-01        27000     2.04       2.07   1081         54009         75
## 6 2010-02-01        29000     1.55       1.58   1700         28901        166
##   treasury_rate closing_price   sigma_20
## 1          0.86         28.41 0.01521071
## 2          0.86         28.41 0.01521071
## 3          0.86         28.41 0.01521071
## 4          0.86         28.41 0.01521071
## 5          0.10         28.41 0.01521071
## 6          0.17         28.41 0.01521071
```

Black Scholes Model: Call and Put Functions

```
black_scholes_put <- function(row){
  S <- as.numeric(row["closing_price"])
  X <- as.numeric(row["strike_price"]) / 1000
  T_ <- as.numeric(row["date_ndiff"]) / 365
  r <- as.numeric(row["treasury_rate"]) / 100
  sigma <- as.numeric(row["sigma_20"])
  d1 <- (log(S / X) + (r  + (sigma ** 2) / 2) * T_) / (sigma * (T_ ** 0.5))
  d2 <- d1 - sigma * (T_ ** 0.5)
  P <- pnorm(-d2) * X * exp(-r * T_) - S * pnorm(-d1)
  P
}

black_scholes <- function(row){
  S <- as.numeric(row["closing_price"])
  X <- as.numeric(row["strike_price"]) / 1000
  T_ <- as.numeric(row["date_ndiff"]) / 365
  r <- as.numeric(row["treasury_rate"]) / 100
  sigma <- as.numeric(row["sigma_20"])
  d1 <- (log(S / X) + (r + (sigma ** 2) / 2) * T_) / (sigma * (T_ ** 0.5))
  d2 <-  d1 - sigma * (T_ ** 0.5)
  C <- S * pnorm(d1) - X * exp(-r * T_) * pnorm(d2)
  C
}
```

Call Options

```
# mean squared error function
mse <- function(df){
  temp <- as.matrix(rowMeans(df[, c("best_bid", "best_offer")])) - as.matrix(df[, c("black_schol
es_pred")])
  squared <- temp ** 2
  sum <- sum(squared)
  mse <- sum / nrow(df)
  mse
}

# median absolute error function
med_abs_err <- function(df){
  temp <- as.matrix(rowMeans(df[, c("best_bid", "best_offer")])) - as.matrix(df[, c("black_schol
es_pred")])
  abs_val <- abs(temp)
  index1 <- nrow(df) %/% 2
  return_val <- abs_val[index1]
  return_val
}

# getting BS prediction values for call options
options_call$black_scholes_pred <- apply(options_call, MARGIN = 1, black_scholes)
head(options_call)
```

```
##         date strike_price best_bid best_offer volume open_interest date_ndiff
## 1 2010-02-01        37500     1.17       1.25    599         41146        719
## 2 2010-02-01        22500     7.20       7.50     22           595        719
## 3 2010-02-01        35000     1.62       1.76      0          2068        719
## 4 2010-02-01        45000     0.40       0.46      0          1697        719
## 5 2010-02-01        27000     2.04       2.07   1081         54009         75
## 6 2010-02-01        29000     1.55       1.58   1700         28901        166
##   treasury_rate closing_price   sigma_20 black_scholes_pred
## 1          0.86         28.41 0.01521071       7.704577e-36
## 2          0.86         28.41 0.01521071       6.287958e+00
## 3          0.86         28.41 0.01521071       1.001357e-20
## 4          0.86         28.41 0.01521071       2.227717e-97
## 5          0.10         28.41 0.01521071       1.415547e+00
## 6          0.17         28.41 0.01521071       3.020968e-03
```

```
call_mse <- mse(options_call)
call_mse
```

```
## [1] 3.336878
```

```
call_med_abs_err <- med_abs_err(options_call)
call_med_abs_err
```

```
## [1] 0.2722713
```

Put Options

```
# getting BS prediction values for put options
options_put$black_scholes_pred <- apply(options_put, MARGIN = 1, black_scholes_put)
head(options_put)
```

```
##         date strike_price best_bid best_offer volume open_interest date_ndiff
## 1 2010-02-01        29000     1.05       1.07   3739         29228         19
## 2 2010-02-01        27000     0.51       0.53   5824          4284         47
## 3 2010-02-01        25000     0.17       0.18    432          1410         47
## 4 2010-02-01        26000     0.29       0.31    351          2364         47
## 5 2010-02-01        28000     0.86       0.89   6838          2956         47
## 6 2010-02-01        26000     0.08       0.09     89          8513         19
##   treasury_rate closing_price   sigma_20 black_scholes_pred
## 1          0.05         28.41 0.01521071       5.892452e-01
## 2          0.10         28.41 0.01521071       6.948123e-23
## 3          0.10         28.41 0.01521071      4.109687e-124
## 4          0.10         28.41 0.01521071       8.027562e-62
## 5          0.10         28.41 0.01521071       1.703454e-04
## 6          0.05         28.41 0.01521071      9.962853e-147
```

```
put_mse <- mse(options_put)
put_mse
```

```
## [1] 5.619199
```

```
put_med_abs_err <- med_abs_err(options_put)
put_med_abs_err
```

```
## [1] 0.93
```

```
# other metrics tested on put options
bid_ask_avg <- as.matrix(rowMeans(options_put[, c("best_bid", "best_offer")]))
bid_ask_avg_minus_pred <- bid_ask_avg - as.matrix(options_put$black_scholes_pred)

rmse <- sqrt(mean(bid_ask_avg_minus_pred ** 2))
rmse
```

```
## [1] 2.370485
```

```
med_err <- median(bid_ask_avg_minus_pred)
med_err
```

```
## [1] 0.34
```

```
avg_abs_err <- mean(abs(bid_ask_avg_minus_pred))
avg_abs_err
```

```
## [1] 1.099471
```

```
med_abs_dev <- median(abs(bid_ask_avg_minus_pred))
med_abs_dev
```

```
## [1] 0.34
```