BACKING OUT THE FUTURE?

A SIMPLE WAY OF LEARNING THE DISTRIBUTION OF
'EXPECTED' FUTURE STOCK PRICES

DAVID
GREENWOOD
DATA SCIENTIST
@DSG22

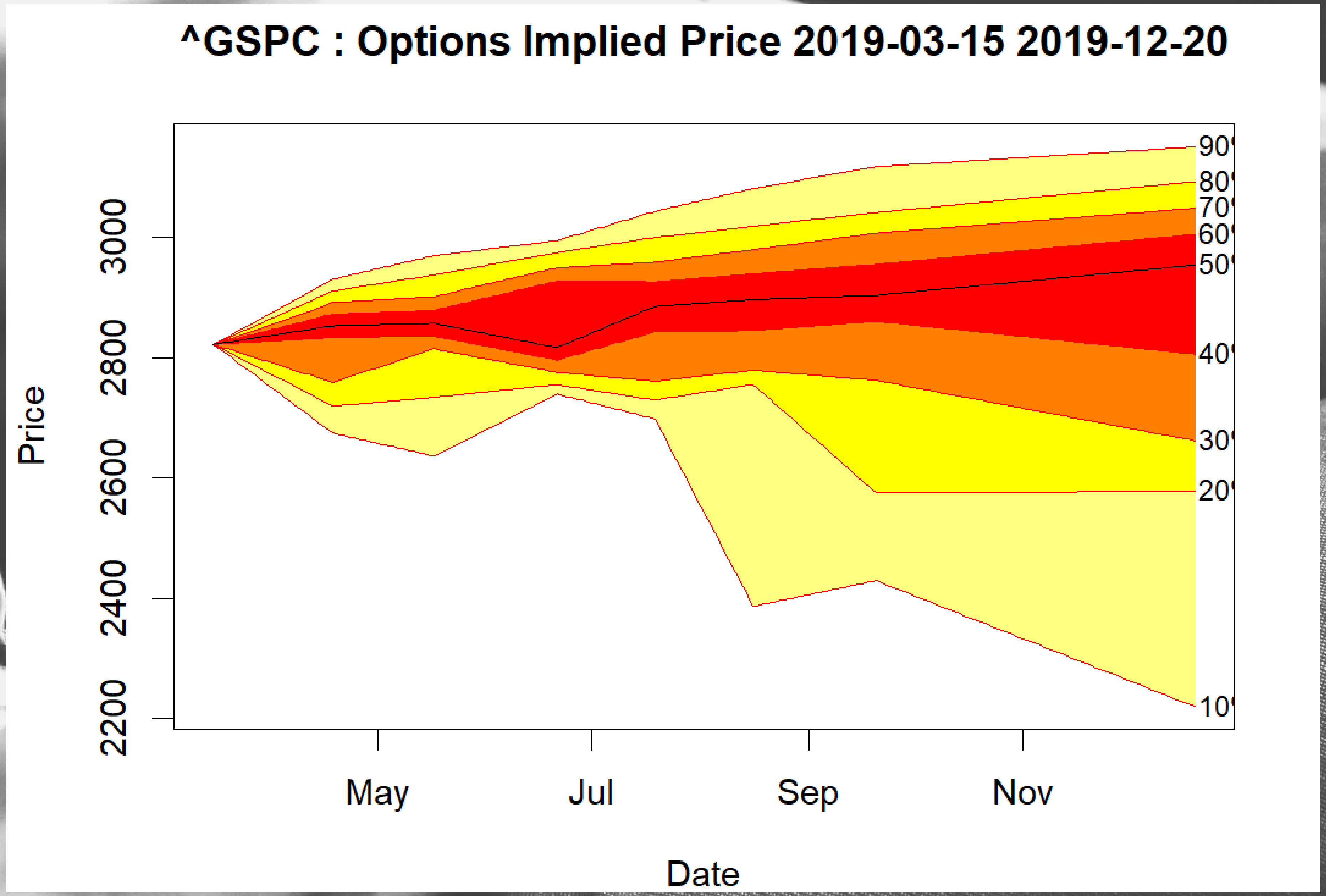
YAHOO
FINANCE
SIDE KICK

MAXIMUM A POSTERIORI
ESTIMATION
DREAM MAKER

CHECK-RISK.COM

LOGICALERRORS.WORDPRESS.COM

WHY INFER 'EXPECTED' FUTURE STOCK PRICES?



CHECK-RISK.COM

LOGICALERRORS.WORDPRESS.COM

LEARNING MARKET IMPLIED INFORMATION

LEARNING CONSENSUS

CHECK-RISK.COM

LOGICALERRORS.WORDPRESS.COM

HEALTH WARNING!

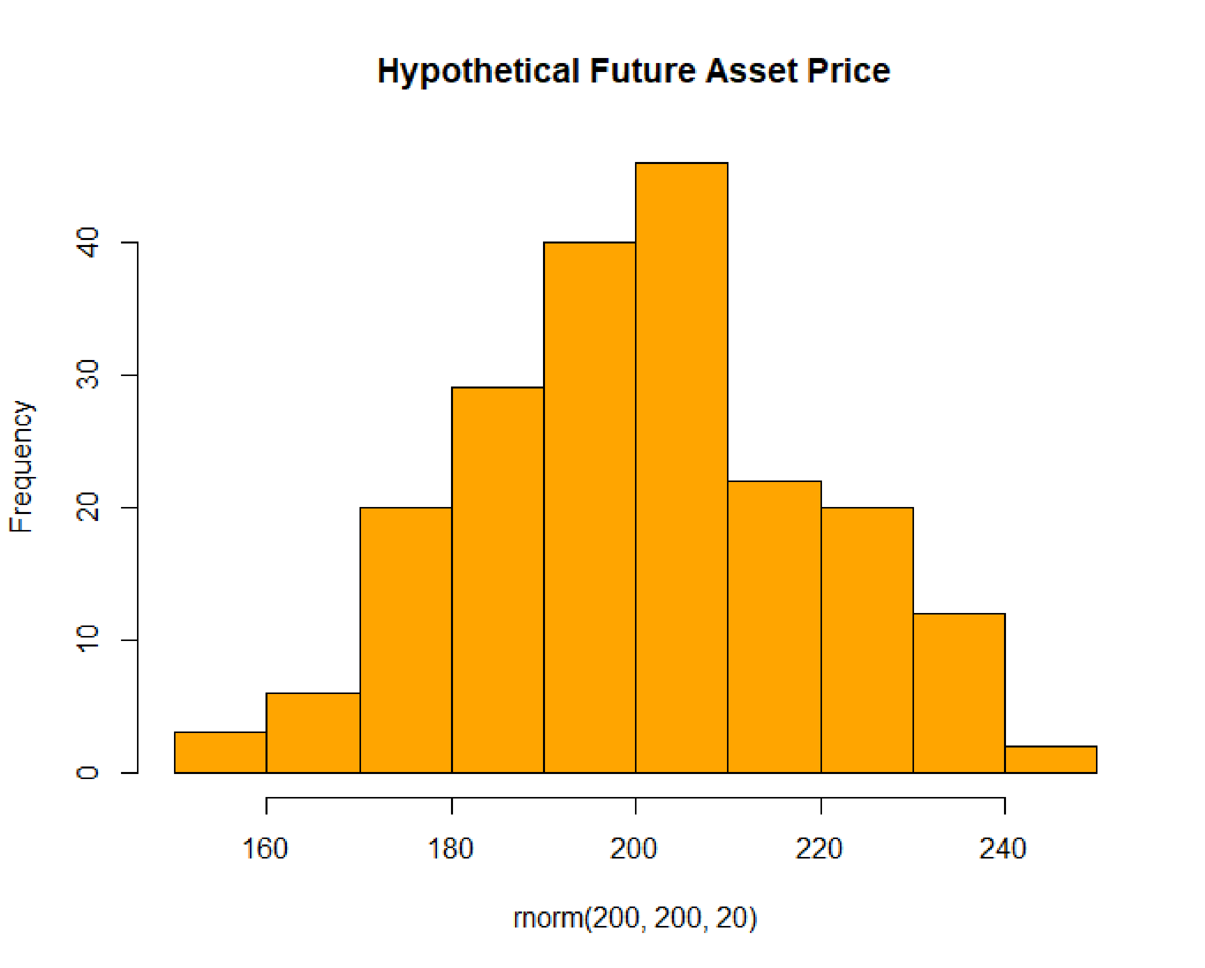
- Market consensus is not necessarily correct!
- The simplifying assumptions used may not always hold!

HOW DOES THIS WORK CONCEPTUALLY?

FINANCIAL MARKETS ARE TIME MACHINES! (OF SORTS)

In theory, one can estimate **future expected prices** by analysing the premia at different "exercise prices" and "exercise dates".

A SIMPLE EXAMPLE



Exercise Date	Exercise Price	Premium	
6 weeks	$ 210.00	$ 14.00	
6 weeks	$ 220.00	$ 10.00	
	Buy 210 & Sell 220	Intrinsic Value	P/L
If SPY under 210 (L)	$ -4.00	$ -	$ -4.00
If SPY 210-220 (M)	$ -4.00	$ 5.00	$ 1.00
If SPY over 220 (H)	$ -4.00	$ 10.00	$ 6.00

CheckRisk

CHECK-RISK.COM

A SIMPLE APPROACH

Exercise Date	Exercise Price	Premium	
6 weeks	$ 210.00	$ 14.00	
6 weeks	$ 220.00	$ 10.00	
	Buy 210 & Sell 220	**Intrinsic Value**	**P/L**
If SPY under 210 (L)	$ -4.00	$ -	**$ -4.00**
If SPY 210-220 (M)	$ -4.00	$ 5.00	**$ 1.00**
If SPY over 220 (H)	$ -4.00	$ 10.00	**$ 6.00**

By assuming that option prices are 'fair' we obtain two simultaneous equations to solve for the probabilities.

$$-4P_L + P_M + 6P_H = 0$$

$$P_L + P_M + P_H = 1$$

$$\frac{5P_L - 1}{5} = P_H$$

$$\frac{6 - 10P_L}{5} = P_M$$


HOW DO YOU
BUILD IT FOR
REAL?
CheckRisk
CHECK-RISK.COM
LOGICALERRORS.WORDPRESS.COM

	A	B	C	D	E	F	G	H
1		Strike	Last	Chg	Bid	Ask	Vol	OI
2	**SPX190418C01000000**	1000	1824.1	0	1825.6	1827.5	5601	23685
3	**SPX190418C01660000**	1660	892.45	0	1163.4	1166.6	1	1
4	**SPX190418C01675000**	1675	877.65	0	1148.4	1151.7	1	1
5	**SPX190418C01700000**	1700	984.6	0	1123.5	1126.7	1	1
6	**SPX190418C01950000**	1950	642.95	0	874.2	877.4	2	2

QUANTMOD IS YOUR FRIEND

1.GET DATA FROM YAHOO FINANCE

```
library(quantmod)
source("./Helper/OptionChainHelper.R")

expiry <- c("2019")

# 1.0 Get data ####
symbol <- "^GSPC" # "^GSPC" "^RUT" "^VIX" "TLT" "VNQ"
stockprice <- getSymbols(symbol, auto.assign = FALSE)
option_chain <- getOptionChain(symbol, Exp = expiry)
```

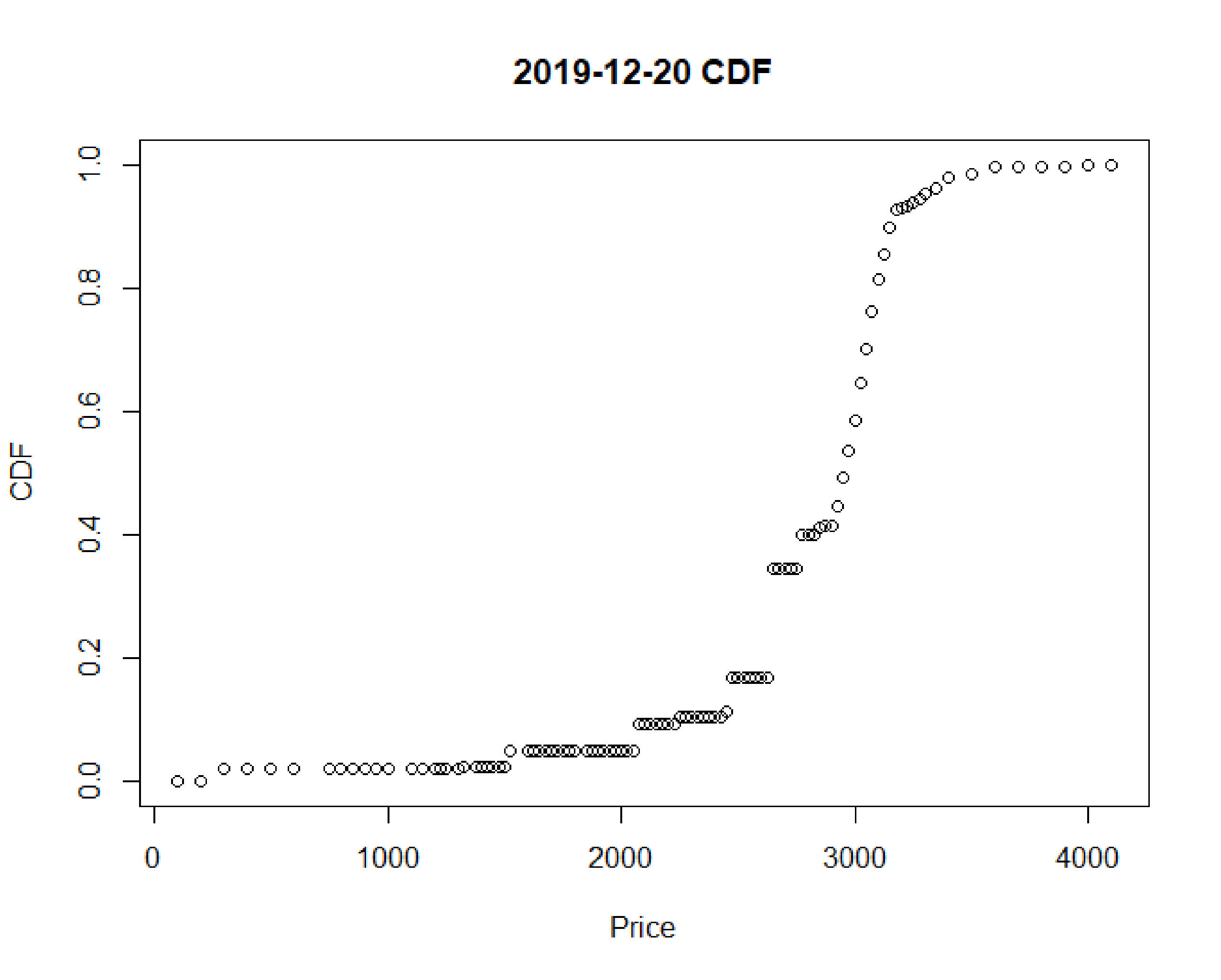
2.CLEAN DATA

```
# 2.0 Pre-process data ####
equity_price <- as.numeric(last(stockprice)[,4]) #get close price
print(paste(symbol, "PX_LAST", equity_price))

processed_chains <- lapply(option_chain, FUN = function(opt) {
  call_chain <- opt$calls
  if (!is.null(call_chain)) {
    call_chain <- subset(call_chain, OI != 0 & Bid != 0 & Ask != 0 ) #filter out data
    call_chain$mid <- (call_chain$Bid + call_chain$Ask)/2
    if (nrow(call_chain) < 10 ) {
      call_chain <- NULL #if insufficient data points then bin
    }
  }
  return(call_chain)
})
names(processed_chains) <- expiry
```

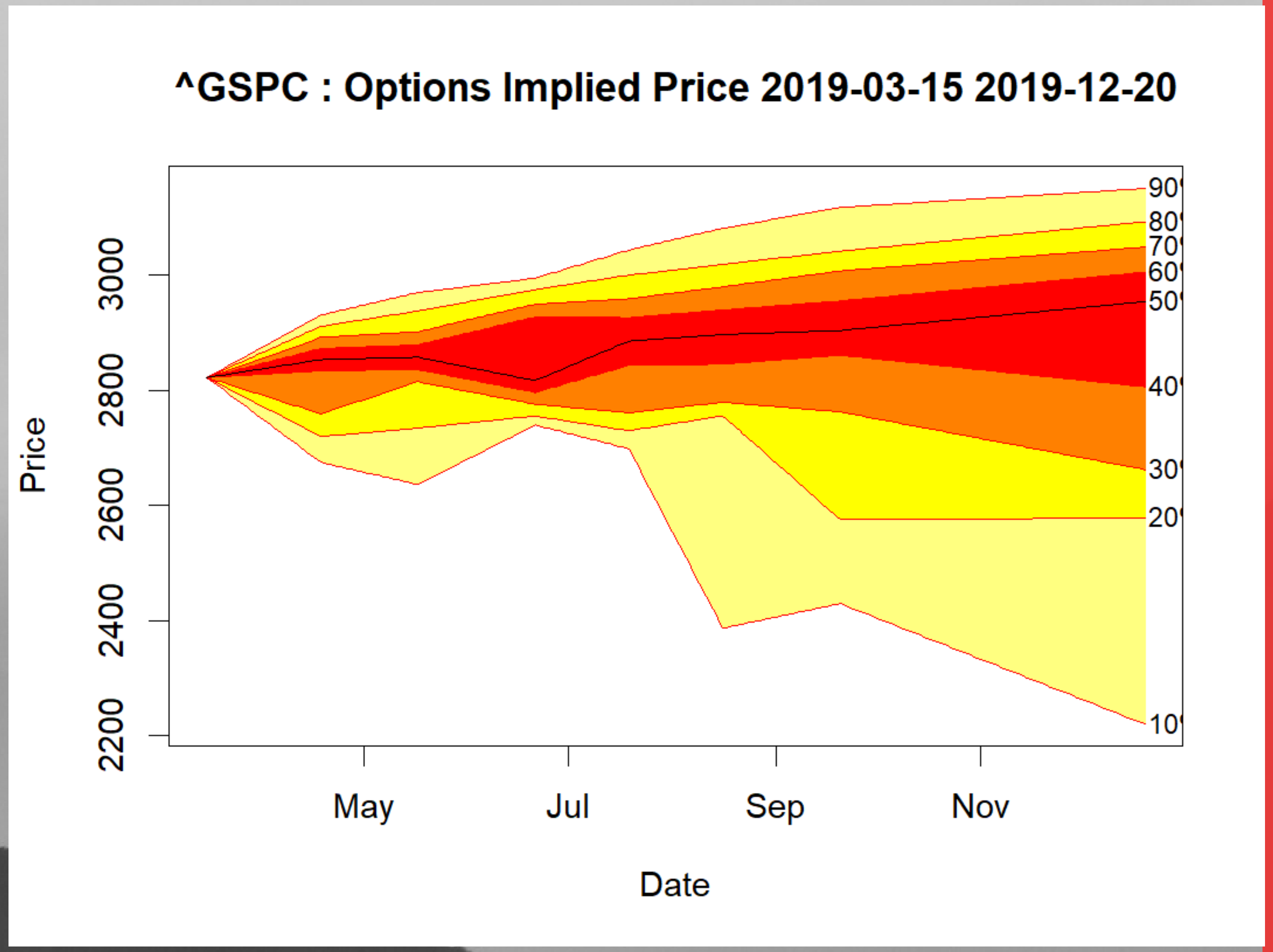
LEARNING FUTURE PRICE PATHS

3. ESTIMATE CDF




```r
# 3.0 Calculate probability ####
# 3.1 #### find empirical cdf
cdfs <- lapply(expiry, FUN = function(e) {
  ps <- NULL
  if (!is.null(processed_chains[[e]])) {
    if (nrow(processed_chains[[e]]) > 0) {
      ps <- estimate_cdf(processed_chains[[e]])
      plot(ps, main = paste(e, "CDF"))
    }
  }
  return(ps)
})
names(cdfs) <- expiry
```

LIBRARY('FANPLOT') IS YOUR FRIEND



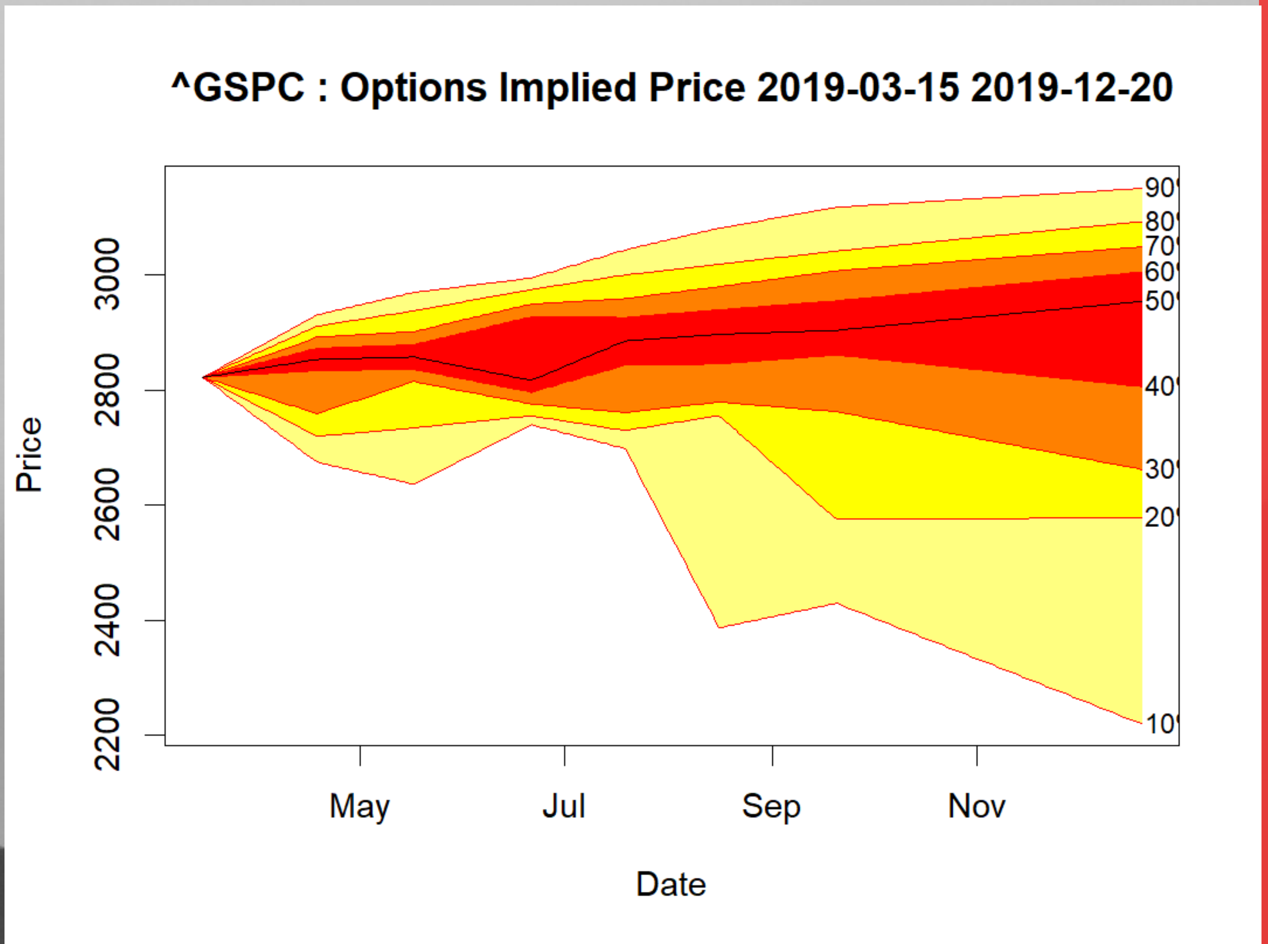

```r
# 5.0 Fan plot ####
library(fanplot)

#5.1 prepare data for fan plot ####
estimate_fan <- preparefan(estimate, days = "all" )#30 "all"

# 5.2 plot fan
plot_fanchart(estimate_fan, p)

plot_fanchart <- function(estimate_fan, p) {
  plot(x = as.Date(names(estimate_fan[5,])),
       y = estimate_fan[5,],
       type = "l",
       ylim = c(floor(min(estimate_fan)), ceiling(max(estimate_fan))),
       main = paste(symbol, ": Options Implied Price",
                    first(colnames(estimate_fan)),
                    last(colnames(estimate_fan))
                    ),
       xlab = "Date",
       ylab = "Price"
  )
  p <- seq(0.1,0.9, 0.1)
  fan(data = estimate_fan, data.type = "values", probs = p,
      start = as.numeric(first(as.Date(names(estimate_fan[5,]))))
      )
  lines(x = as.Date(names(estimate_fan[5,])), y = estimate_fan[5,])
}
```

CHECK-RISK.COM



SUMMARY

YOU CAN LEARN THE DISTRIBUTION OF 'EXPECTED' STOCK PRICES BY:

1.GETTING OPTION CHAIN DATA FROM YAHOO FINANCE - LIB(QUANTMOD)
2.CLEANING OUT MISSING DATA
3.CREATING CDF FOR EACH EXPIRY
4.CREATE FAN PLOT - LIB('FANPLOT')

BACKING OUT THE FUTURE?

A SIMPLE WAY OF INFERRING THE DISTRIBUTION OF
'EXPECTED' FUTURE STOCK PRICES

DAVID
GREENWOOD
DATA SCIENTIST

YAHOO
FINANCE
SIDE KICK

MAXIMUM A POSTERIORI
ESTIMATION
DREAM MAKER

CHECK-RISK.COM

A SIMPLE APPROACH

How do we reasonaly know Prob of L?

	Strike	Last	Chg	Bid	Ask	Vol	OI
SPX190418C01000000	1000	1824.1	0	1825.6	1827.5	5601	23685
SPX190418C01660000	1660	892.45	0	1163.4	1166.6	1	1
SPX190418C01675000	1675	877.65	0	1148.4	1151.7	1	1
SPX190418C01700000	1700	984.6	0	1123.5	1126.7	1	1
SPX190418C01950000	1950	642.95	0	874.2	877.4	2	2
SPX190418C02000000	2000	827.85	7.849976	827.9	829.9	8	23700
SPX190418C02050000	2050	685.38	0	774.5	777.8	1	4
SPX190418C02100000	2100	505.4	0	724.7	727.9	1	5
SPX190418C02120000	2120	348.1	0	704.8	708	0	0
SPX190418C02140000	2140	281.7	0	684.9	688.1	0	10
SPX190418C02150000	2150	379.7	0	674.9	678.1	0	20
SPX190418C02200000	2200	627.37	4.969971	628.6	630.6	2	2970
SPX190418C02250000	2250	578.8	0	578.5	580.6	2	42
SPX190418C02260000	2260	215.5	0	565.5	568.7	0	0
SPX190418C02270000	2270	367.7	0	555.5	558.7	6	3
SPX190418C02280000	2280	309.5	0	545.6	548.8	1	10
SPX190418C02300000	2300	510.5	0	529.3	531.4	1	174
SPX190418C02325000	2325	485.65	0	504.5	506.5	1	22
SPX190418C02350000	2350	270.7	0	476	479.2	2	37
SPX190418C02360000	2360	256	0	466.1	469.3	0	15
SPX190418C02365000	2365	253.82	0	461.1	464.3	1	2
SPX190418C02375000	2375	256.6	0	454.5	456.6	10	24