

CS 7646 Exam 1
October 12, 2017
Exam Version A

***Do not open this
booklet until
instructed to begin***

1. What is the output of the following Python session?

```
>>> def func(x = 2):  
...   b = 2  
...   return x * b  
...  
>>> x = 3  
>>> b = 3  
>>> print func()
```

- A) 2
- B) 6
- C) 4
- D) 9

2. What is the output of the following Python session?

```
>>> a = 3  
>>> b = a  
>>> a = 2  
>>> print b * a
```

- A) 4
- B) 6
- C) 3
- D) 2

3. Consider the following Python session and its output. What should you replace XXXX with in order to get the following output?

```
>>> a = np.random.uniform(size=(3,2))  
>>> b = a/a[1,:]   
>>> print XXXX  
1.0
```

- A) b[0,1]
- B) b[1,1]
- C) a[0,1]
- D) a[1,1]

4. Consider the following data:

COL_1	COL_2	COL_3
100.00	0.00	0.00
101.00	0.01	0.01
102.00	0.02	0.01
100.00	0.00	-0.02
102.00	0.02	0.02

What might the data in each of the columns represent (from left to right)?

- A) Cumulative return, portfolio value, daily return
- B) Daily return, portfolio value, cumulative return
- C) Cumulative return, daily return, portfolio value
- D) Portfolio value, cumulative return, daily return

5. Which code snippet below would correctly calculate the Sharpe ratio for time series `daily_rets` that represents 60 days of daily returns? The risk free rate is represented by `rfr`

```
# snippet A
sharpe = math.sqrt(60.0) * np.mean(daily_rets - rfr) / np.std(daily_rets)
```

```
# snippet B
sharpe = math.sqrt(60.0) * np.std(daily_rets - rfr) / np.mean(daily_rets)
```

```
# snippet C
sharpe = math.sqrt(252.0) * np.mean(daily_rets - rfr) / np.std(daily_rets)
```

```
# snippet D
sharpe = math.sqrt(252.0) * np.std(daily_rets - rfr) / np.mean(daily_rets)
```

6. Consider the following code snippet. What code could you replace with `XXXX` to cause the following output?

```
>>> a = np.random.uniform(size=(3,2))
>>> a
array([[ 0.9598058 ,  0.30898308],
       [ 0.08671576,  0.83212644],
       [ 0.9701383 ,  0.25618103]])
>>> b = np.random.uniform(size=(3,2))
>>> b
array([[ 0.363649 ,  0.86394158],
       [ 0.89838762,  0.46682958],
       [ 0.76027507,  0.46149348]])
>>> XXXX
>>> a
array([[ -1.          ,  0.30898308],
       [ 0.08671576, -1.          ],
       [ 0.9701383 , -1.          ]])
```

- A) `a[0,0] = -1`
- B) `a[a<0.5] = -1`
- C) `a = a/a[0,:]`
- D) `a[b<0.5] = -1`

7. Which statement is TRUE?

- A) Mutual funds can be traded throughout the trading day.
- B) Hedge fund holdings are fully transparent.
- C) ETF managers earn 20% of the ETF's profits.
- D) ETFs can be traded throughout the trading day.

8. Which statement is FALSE?

- A) Hedge fund managers typically charge 2% of profits and 20% of AUM as fees.
- B) Mutual fund managers typically charge 0.25% to 2.0% of AUM as fees.
- C) ETF managers typically charge 0.1% to 1.0% of AUM as fees.
- D) Hedge fund managers typically charge 2% of AUM and 20% of profits as fees.

9. Consider two hedge funds. Both of them have provided equivalent positive cumulative returns. Which of the following might be a valid reason for choosing HF1 over HF2?

- A) HF1 has a lower Sharpe ratio than its benchmark, but HF2 has a higher Sharpe than its benchmark.
- B) HF1 has a higher standard deviation of daily returns than HF2.
- C) HF1 has a lower standard deviation of daily returns than HF2.
- D) HF1 has a higher Bollinger value than HF2.

10. According to the 60 minutes video "Is the stock market rigged?" which strategies have high frequency traders used to exploit market mechanics?

- A) A dedicated fiber optic link between Silicon Valley and the exchanges in New York.
- B) A dedicated fiber optic link between Chicago and the exchanges in New Jersey.
- C) Front running orders by exploiting faster network connectivity.
- D) B and C.

11. How does the IEX exchange defeat the high frequency traders investigated in the 60 minutes video?

- A) Using dark pools.
- B) Using 60 kilometers of fiber optic link to delay access to the order book.
- C) By converting the market to a series of discrete auctions every 2 seconds.
- D) By prohibiting limit orders.

12. Consider the following order book. At what average price would a limit order to BUY 100 shares at 100.05 be executed at?

Bid size	Price	Ask size
	100.10	300
	100.05	300
	100.00	200
100	99.95	
50	99.90	
100	99.85	

- A) 100.00
- B) 100.025
- C) 100.05
- D) 100.10

13. Consider the following order book. At what average price would a market order to SELL 250 shares be executed at?

Bid size	Price	Ask size
	100.10	300
	100.05	300
	100.00	200
100	99.95	
50	99.90	
100	99.85	

- A) 99,95
- B) 99.90
- C) 99.85
- D) None of the above

14. Consider the following order book. How would you expect the price to change over the next few minutes?

Bid size	Price	Ask size
	100.10	100
	100.05	100
	100.00	200
1000	99.95	
500	99.90	
500	99.85	

- A) The price will go up because there is buying pressure.
- B) The price will go down because there is selling pressure.
- C) The price will stay about the same.
- D) The price will remain fixed at 99.95

15. Consider the following valuation factors of a company:

It owns 1000 cars valued at \$20,000 each
It holds patents worth \$5,000,000
It owes \$5,000,000 in loans
It pays \$1.00 per year per share in dividends starting in one year
The stock price is \$60.00 per share
There are 1,000,000 shares outstanding
The discount rate is 2%

What is the book value of the company?

- A) \$25,000,000
- B) \$20,000,000
- C) \$15,000,000
- D) \$30,000,000

16. Consider the following valuation factors of a company:

It owns 1000 cars valued at \$20,000 each
It holds patents worth \$5,000,000
It owes \$5,000,000 in loans
It pays \$1.00 per year per share in dividends starting in one year
The stock price is \$60.00 per share
There are 1,000,000 shares outstanding
The discount rate is 2%

What is the intrinsic value of the company?

- A) \$25,000,000
- B) \$20,000,000
- C) \$15,000,000
- D) \$50,000,000

17. Consider the following valuation factors of a company:

It owns 1000 cars valued at \$20,000 each
It holds patents worth \$5,000,000
It owes \$5,000,000 in loans
It pays \$1.00 per year per share in dividends starting in one year
The stock price is \$60.00 per share
There are 1,000,000 shares outstanding
The discount rate is 2%

Assume the price of the stock should settle to book value + 10%. What would you expect to happen to the price in the future?

- A) The price should stay about the same.
- B) The price should go up.
- C) The price should go down.
- D) A shrubbery.

18. According to the CAPM, which statement(s) are true regarding Alpha and Beta? Assume the market has gone up 1% in the last day.

- A) If a stock's Alpha is 2% and the Beta is 1.0, the stock should go up 2%
- B) If a stock's Alpha is 1% and the Beta is 2.0, the stock should go up 2%
- C) If a stock's Alpha is 1% and the Beta is 1.0, the stock should go up 2%
- D) If a stock's Alpha is 2% and the Beta is 2.0, the stock should go up 3%

19. According to the CAPM, which statement(s) are true regarding Alpha and Beta?

- A) The expected value of Beta is 0.0.
- B) The expected value of Alpha is 0.0.
- C) We expect Beta = Alpha
- D) We expect Alpha > Beta

20. If the weak form of the EMH is true, which sorts of trading advantage(s) would be eliminated?
- A) Insider information only.
 - B) Fundamental information only.
 - C) Technical information only.
 - D) All advantages including: Insider, Fundamental and Technical.
21. Suppose a stock has split (1 share becomes 2 shares) 4 times between 2010 and 2016 and that you use a historical price service that adjusts prices for splits and dividends. If the actual market close price for the stock on July 1, 2009 was \$100.00 and on July 1 2017 it was \$200, what would you expect the adjusted close to be on those two dates respectively?
- A) \$6.25 in July 2009, and \$200.00 in July 2017.
 - B) \$100.00 in July 2009, and \$3200.00 in July 2017.
 - C) \$100.00 in July 2009, and \$200.00 in July 2017.
 - D) \$400.00 in July 2009, and \$800.00 in July 2017.
22. Suppose you have historical stock price data with data missing on some days in history (the values are NaN). You still want to use the data in backtesting and calculation of technical factors. Which of the following options are recommended (in the book)?
- A) Replace the NaNs with zeros.
 - B) Fill the data back only.
 - C) Replace the NaNs with the average price.
 - D) Fill forward, then fill back.
23. Consider kNN, linear regression (LR), and Decision tree (DT) learning (using correlation for splitting). Which option correctly lists the methods from fastest to slowest in training time?
- A) kNN, LR, DT
 - B) LR, kNN, DT
 - C) DT, LR, kNN
 - D) LR, DT, kNN
24. Consider kNN, linear regression (LR), and Decision tree (DT) learning (using correlation for splitting). Which option correctly lists the methods from fastest to slowest in query time?
- A) kNN, LR, DT
 - B) LR, kNN, DT
 - C) DT, LR, kNN
 - D) LR, DT, kNN
25. Consider two decision trees trained on the exact same data. DT was trained using correlation for splitting, RT was trained using splits determined randomly. Both trees were trained with leaf_size = 1. Which option below correctly describes (in order): The fastest to train, the fastest to query, the highest accuracy on in-sample data?
- A) DT, DT, DT
 - B) RT, about the same, about the same
 - C) RT, RT, about the same
 - D) DT, RT, about the same

26. Consider overfitting when using kNN and decision trees. When overfitting occurs with these two methods, in which “direction” does it occur?

- A) As k increases; As leaf_size increases
- B) As k increases; As leaf_size decreases
- C) As k decreases; As leaf_size decreases
- D) As k decreases; As leaf_size increases

27. Which of the following is NOT an advantage of ensemble learners?

- A) Less susceptibility to overfitting.
- B) Reduction in RMSE when using weak learners.
- C) Improved computational speed compared to individual weak learners.
- D) Can work with many types of weak learners.

28. Which of the following is NOT typically used as a means of assessing a learning algorithm?

- A) The standard deviation of predicted values.
- B) Compute time for training.
- C) Compute time for querying.
- D) Correlation of predicted values with known correct values.

29. Consider a data set composed of 1000 samples where X is drawn randomly uniformly from -2π to $+2\pi$, and $Y = \sin(X)$ (two full sine wave cycles). Consider kNN, decision trees, random trees and linear regression. Which statement is true regarding in-sample RMSE?

- A) Linear regression will perform best.
- B) kNN, decision trees and random trees will all do better than linear regression.
- C) kNN and linear regression will perform about the same.
- D) Decision trees will perform significantly better than random trees.

30. Suppose you are using one of the minimizing optimizers from ScikitLearn. You are using it to optimize your portfolio for MAXIMUM cumulative return, and port_vals are the daily total values of the portfolio for a particular allocation. Which of the following would be the best way to compute the objective function for the optimizer?

- A) `objective = port_vals.max()`
- B) `objective = - port_vals.max()`
- C) `objective = port_vals[-1]/port_vals[0]`
- D) `objective = - port_vals[-1]/port_vals[0]`