DS 610 Week 8 Assignment
Big Data Analytics

Due Date:

Please prefix the file name with your first name : **name_ds610_assignment_08**

1. Write a python program which returns the reverse of a number without using string operations. Explain your logic.
Sample Input : 1234
Sample Output : 4321

2. What is the name of Hive's SQL implementation?
A) HiveQL
B) SQL on Hive
C) HiveStructured
D) Hive SQL

3. Tables in Hive have a schema
A) True
B) False

4. What type of tables are stored in the Hive Data warehouse?
A) Separate
B) Internal
C) External
D) Managed

5. What is stored in Hive for an External Table?
A) Create table syntax
B) Definition
C) Copies of the data
D) Data

6. What type of partitions are typically used to shrink down table size?
A) Date
B) Location
C) Categorical
D) Integer

7. Grouping sets in Hive is a shortcut for performing what types of operations?
A) Aggregations
B) Calculations
C) Summarization
D) Filtering

8. In Hive, what does the CUBE function return?
A) All possible aggregate combinations
B) Hierarchical aggregate combinations
C) Group by summarization
D) Standard deviation analysis

9. When does the HAVING clause apply its filter to your result set?
A) After the order by is complete
B) After aggregations are complete
C) Before aggregations are complete
D) With a group by

10. Which filtering clause allows for a wildcard?
A) LIKE
B) WILDCARD
C) REGULAR
D) EXP()

11. Which join type only returns data from two tables where the join condition is true?
A) Inner join
B) Wildcard join
C) Outer join
D) Left join

12. What type of Hive function allows for IF statements?
A) Conditional
B) Date
C) String
D) Math

13. Which function will return the number of days between two dates?
A) datediff
B) days_between_dates
C) duration
D) datesiff

14. Which function combines multiple strings into a single value?
A) combine_string
B) combine
C) concat_string
D) concat

15. Using *sales.csv*, write a HiveQL to return the number of rows in target table. Submit the screenshot of your Hive view showing query and the result.

16. Using *clients.csv*, write a HiveQL to return the number of rows in target table. Submit the screenshot of your Hive view showing query and the result.

Thank you.