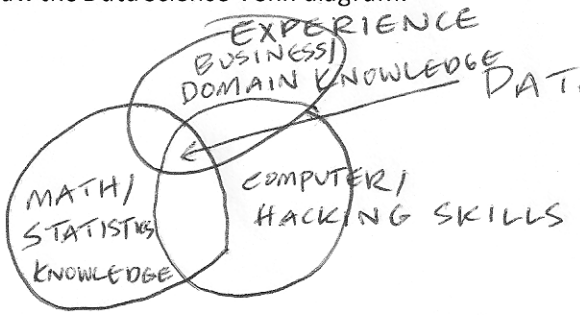
KEY

DNSC 6279 – Quiz 1

1. **(4 pts.)** Draw the Data Science Venn diagram:



(ORDER IN CIRCLES DOES NOT MATTER)

(Include at least the three major labels and the location of Data Science in the diagram.)

2. **(1 pt.)** Give an example of. or describe, structured data.

NEAT ROWS AND COLUMNS OF NUMERIC DATA, BILLING INFORMATION
DATA ORGANIZED INTO MATRIX, TRANSACTIONS
CUSTOMER DATA, DEMOGRAPHIC DATA

3. **(1 pt.)** Give an example of, or describe, unstructured data.

DATA LIKE TEXT OR IMAGES THAT IS NOT ARRANGED INTO COLUMNS AND ROWS (LIKE TWEETS, IMAGES, SOUND, VIDEO) OR DATA WITHOUT SCHEMA, LOGS

4. **(1 pt.)** What is the basic difference between supervised learning and unsupervised learning?

IN SUPERVISED LEARNING, THE "Y"/DEPENDENT/TARGET/LABEL VARIABLE IS KNOWN

5. **(1 pt.)** Give an example of, or describe, deploying a predictive model.

MOVING A PREDICTIVE MODEL FROM A DEVELOPMENT ENVIRONMENT INTO AN OPERATIONAL/PRODUCTION ENVIRONMENT.
EXAMPLES: INSURANCE PRICING, CREDIT CARD TRANSACTIONS, PRODUCT RECOMMENDATION

6. **(1 pt.)** Give an example of a legitimate business decision process in which the use of analytics would be inappropriate or superfluous?

MODELING THE EXTREMELY UNLIKELY & UNIMPORTANT
MODELING WHEN DOMAIN EXPERTISE SHOULD BE USED INSTEAD
CONFIRMING THE OBVIOUS
WHEN NO PAST DATA IS AVAILABLE

7. **(1 pt.)** What is more important in data science: predictive modeling OR making inferences about past events?

NEITHER ARE MORE IMPORTANT

BOTH ARE IMPORTANT

IT DEPENDS ON THE SITUATION