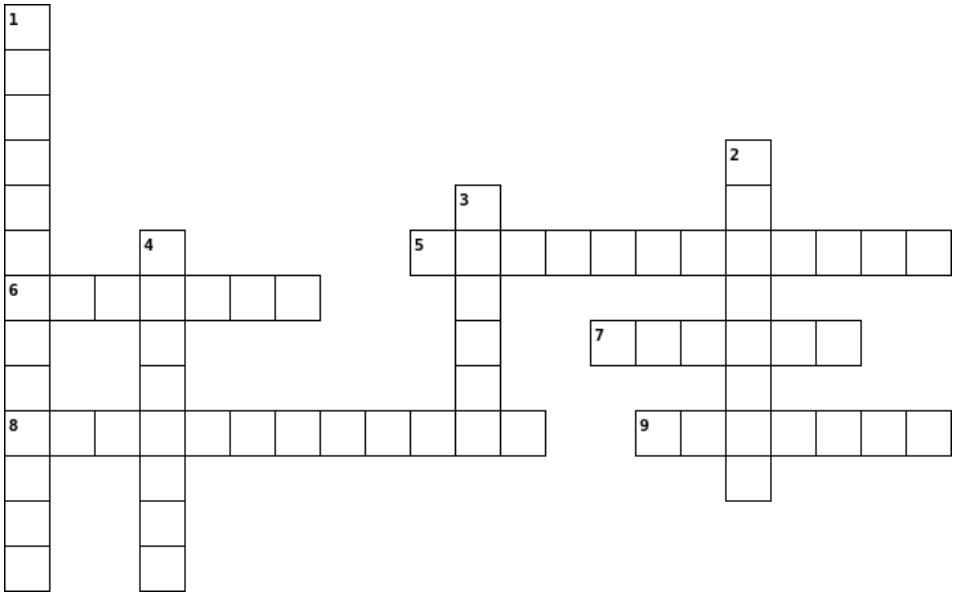
Data Science Quiz

Question 1

Zoology and Botany in Data Science *(hint:ignore whitespaces in words)* *(9 points)*



Down

1. Are prone to overfitting polynomials

2. Has a memory like Postgres

3. Have six fingers

4. Friendly marine mammal that carries containers Docker/whale

Across

5. Microblogging mountain bluebird named after basketball legend tweeter

6. Feline with 8 branches

7. Nonvenomous ambush predator

8. If you spill a basket of acorns, you get this

9. Data management the natural way

Question 2

Find 10 bugs: *(10 points)*

```
for year in range(1890, 2015, 1):
    total = 0
    filename = 'names\yob{year}.txt'
    for line in open(filename, 'w'):
        columns == line.split(',').strip()
total =+ columns[2]

print("Result: {} births total".format(year))
```

```
total = 0
for year in range(1890, 1891, 1):
    filename = 'names\yob'+str(year)+'.txt'
    for line in filename:
        columns = line.strip().split()
        total += columns[0]
print(f"Result: {total} births total")
```

Question 3

What do the following bash commands do? *(6 points)*

`ls -a`	reveal files + hidden files
`sudo rm -rf /`	
`chmod 700 *`	
`grep print *.py | wc -l`	

Question 4

Write an SQL query that extracts the 10 most frequently occurring items in the *'subject'* column from the table *'data_scientists'*, but only consider students with the column *'python'* being 1 or higher. Output results in descending order. *(10 points)*

Question 5

Name the functions. *(9 points)*

$P(A\|B)=\frac{P(B\|A)P(A)}{P(B)}$	
$\frac{1}{N}\sum_i (y_i - y_i^{true})^2 + \lambda \sum_j b_j^2$	
$\frac{1}{1+e^{-x}}$	

Question 6

Write 2 items you could import from each Python module. *(10 points)*

`pandas`	`DataFrame`
`random`	
`numpy`	
`seaborn`	
`os`	

Question 7

Which strings does the Regular Expression `'R[oau]\w+e'` match? *(8 points)*

Rome	rose	Rue	Dome
Rhizome	Rhizome	Ru\w+e	Raave

Question 8

Write the Docker one-liners to do the following *(10 points)*

Run and start a standard python container with the name `my_python`	
Interact with `my_python`	
Calculate 4+4 in `my_python`	
Stop `my_python`	
Delete `my_python`	

Question 9

Match pairs. *(7 points)*

	postgreSQL		MongoDB
1	`Table`	1	`use sales`
2	`Row`	2	`show collections`
3	`\l`	3	`Document`
4	`\dt`	4	`db.sales.find()`
5	`\c sales`	5	`Collection`
6	`SELECT * FROM sales;`	6	`db.sales.distinct('client')`
7	`SELECT DISTINCT client FROM sales;`	7	`show dbs`