

Name: _____

MAT 339 Data Science Techniques - - Quiz 1

11 February 2026

20 points

Time: 20–25 minutes

You may use one 3" × 5" note card hand-written and/or typed by you, notes on both sides allowed.
All questions refer to commands that can be executed in Git Bash or Anaconda PowerShell unless otherwise noted.

1. (3 points) Briefly describe what each of the following commands accomplishes.

(a) `ls -l`

(b) `echo "hello world" | tr [j-t] [J-T]`

(c) `grep error logfile.txt | tee errors.txt`

2. (3 points) Write a command (or commands) covered in class and/or the homework to accomplish each of the following tasks (there may be more than one way to accomplish each):

(a) download several files with a single command; assume the files are located at `http://www.us.gov/`, have the names `data01.dat`, `data02.dat`, `data03.dat`, `data04.dat` and that we want the files to have the same names once downloaded

(b) count the number of lines in a file named `system-log.txt`

(c) view the first 4 lines of a file named `speech.md`

3. (1 point) Which command prints the names of all markdown (.md) files within the current directory and its subdirectories?

- A. `ls *.md`
- B. `find . -type d -name "*.md"`
- C. `find . -name "*.md"`
- D. `grep ".md" *`
- E. `find *.md`

4. (1 point) Suppose `data01.csv` is a `csv` file. Which command prints the **unique values from the 8th column**, sorted alphabetically?

- A. `cut -f8 data01.csv | uniq | sort`
- B. `cut -d',' data01.csv | uniq | sort`
- C. `cut -d',' -f8 data01.csv | uniq | sort`
- D. `cut -d',' -f8 data01.csv | sort | uniq`
- E. `csvcut -c 8 data01.csv | uniq`

5. (3 points) Indicate whether each statement is **True** or **False**.

- | | | |
|--|------|-------|
| (a) If two collaborators on a GitHub repo make overlapping changes to a file, <code>git</code> can automatically resolve the conflict. | True | False |
| (b) The <code>\D</code> regular expression shorthand character class signifies match a single digit. | True | False |
| (c) The command <code>csvgrep -c 3 -m 42 data01.csv</code> searches for the value 42 anywhere in the third column of <code>data01.csv</code> . | True | False |

6. (4 points) Match the Git command on the left with its purpose on the right. Write the letter of the correct purpose in the blank next to each command. Some options will not be used.

Git Command	Purpose
1. <code>git pull</code> ____	A. Record a snapshot of changes in your local repository
2. <code>git commit</code> ____	B. Stage changes to be included in the next commit
3. <code>git push</code> ____	C. Download a repository from a remote server
4. <code>git clone</code> ____	D. Retrieve changes from a remote repository and merge locally
	E. Upload committed changes to a remote repository
	F. Show the status of files

7. (2 points) What does `csvstat data.csv` provide that `csvlook data.csv` does not?

8. (2 points) Which `csvkit` command allows you to select columns by name or number from a `csv` file?

9. (1 point) In a complete sentence or two, explain one or more advantages of using the terminal over alternatives.