

Khan Academy Responses

```
<!DOCTYPE html>
<html>
<head>
<title>Khan Academy Responses</title>
</head>
<body>

<h1>Khan Academy Responses</h1>
<h3><b><u>Data Tools</u></b></h3>
<p>This part wasn't too surprising to me since I am in an information
foundations class right now, and I had to use Excel a ton. Datasheets tend
to be something a lot of companies use since they are very easy to
navigate for users. CSV's are harder to read but are the basic format for
many tools. The flaw is the fact that it holds only one file, and often
times once the file gets larger, the processing time increases. One thing
that helps a lot is the ability to upload and change CSV sheets into
spreadsheets and database formats and save them as a CSV later, that way
it can be easily read and edited by most people.</p>
<h3><b><u>Big Data</u></b></h3>
<p>I knew there was a lot of data within the world of the interweb, but I
never knew the exact number. 175 zettabytes is a much larger number than
most humans can comprehend. My theory about what file sizes were bigger
was correct too, from smallest to largest, text, image, gif, audio, video,
then audio+video files. Meaning, keeping things as image and text files
only, at least as much as possible, is the best way to save on data space.
It was also fun seeing that these "data wear houses," are just large
computer rooms with shelves upon shelves of hard-disc drives. </p>
<p>One last little side note Facebook's data output is very unsurprising to
me. From all I have read and seen, that app is one of the largest data
mines to date.</p>
<h3><b><u>Bias in Machine Learning</u></b></h3>
<p>Reading about this in hello world, and seeing it in my criminal justice
class, nothing was new about this. AI and machine learning programs learn
from human input, and depending on the sources, can turn up very
undesirable results. Supervised or unsupervised, most machine learning
will learn human bias. Blacks being flagged more than whites being the
prime example. Face recognition seeing people who look similar and causing
innocent people to get arrested. It's just a big mess.</p>
<h3><b><u>Unit Test</u></b></h3>
```

<p>I don't really like tests of any form, but this one wasn't too stressful. If you got an answer wrong, it explained it to you, and it explained why the answer was right, which I did enjoy. It didn't really feel like a test, just a review. I wish there were more activities like this instead of high-stress one-chance tests. </p>

</body>

</html>

Questions, Hypotheses, and Data

```
<!DOCTYPE html>
<html>
<head>
<title>Questions, Hypotheses, and Data</title>
</head>
<body>

<h1>Questions, Hypotheses, and Data</h1>
<h3><b><u>Questions</u></b></h3>
<p><b>1.</b>How many users worldwide does Facebook have?
</p>
<p><b>2.</b>Which Country has the most users?
</p>
<p><b>3.</b>Which country has the most users per country population?
</p>

<h3><b><u>Hypotheses</u></b></h3>
<p><b>-</b>There are over 1 million users per country, or at least for
most countries.
</p>
<p><b>-</b>Being an American-based company, I would guess the US has the
most users.
</p>
<p><b>-</b>America probably has the highest percentage as well.
</p>
<p><b>-</b>Countries with fewer electronics (such as phones and computers)
will have fewer users.
</p>

<h3><b><u>Data Links</u></b></h3>
<p> Data Link from Data Base:
<a
herf="https://www.kaggle.com/datasets/faisaljanjua0555/facebook-users-by-c
ountry-data-cleaned?resource=download&select=data.csv/">FaceBook Users By
Country Data (Cleaned)</a>
</p>
<p> Data Link Saved to GoogleSheets:
```

```
<a
herf="https://docs.google.com/spreadsheets/d/1Hn5pIkgpWuNuiP2TdHdLUF5V7qLz
3ik2cwPbmu_Sp8A/edit?usp=sharing/">FaceBooker Users By Country
</a></p>
```

```
</body>
```

```
</html>
```

Process and Analysis

```
<!DOCTYPE html>
<html>
<head>
<title>Process and Analysis</title>
</head>
<body>

<h1><b>Process and Analysis</b></h1>

<h3><u>Process</u></h3>
<p>My "problem" was more of a curiosity. So many people I know have a
Facebook, and many of them only have a Facebook, and only use it. I got my
data set from Kaggle by searching "Facebook users by country," and it
pulled up a table with over 200 countries. Some of the data I got was from
simple Google searches since it's all numerical data I am looking for, but
most of it I got from the table.</p>

<h3><u>Analysis</u></h3>
<p>My analysis was rather straightforward. Using things such as google
sheets, I can use sum functions, charts, and filters (such as max and
min), to answer my questions and prove/disprove my hypothesis. Things the
data couldn't answer, a simple Google search was how I got the answer.
</p>

<h3><b><u>Questions & Answers</u></b></h3>
<p><b>1.</b>How many users worldwide does Facebook have?
</p>
<p>2.96 Billion, which is a lot more than I was expecting. That's about
36.7% of the population of the world, which is a massive amount when it
comes to social media platforms. Especially one known for such terrible
data leaks. </p>
<p><b>2.</b>Which Country has the most users?
</p>
<p>Despite my hypothesis, India has the most users, but I guess I should
have guessed that since they have almost 4 times the US Population. </p>
<p></p>
<p><b>3.</b>Which country has the most users per country population?
</p>
```

<p>Surprisingly, based on the data, none of the countries with the highest user count don't have very high percentages.

Vatican City had the highest percentage, which leads me to believe there are either over 100 business accounts in the country, or there is a miscount in the data.

</p>

<p></p>

<p>While I was getting this data set, I accidentally put the filter to min-max, which showed me North Korea had the smallest user population percentage, which, was not surprising. At 0.05% population to user percentage, the most surprising number was the fact North Korea had 14K users, since that country has the most strict internet laws.</p>

<h3><u>Hypotheses Answers</u></h3>

<p>-There are over 1 million users per country, or at least for most countries.

</p>

<p>This was only partially correct, most countries did indeed have over 1M users, but a lot more than I was expecting had fewer. </p>

<p>-Being an American-based company, I would guess the US has the most users.

</p>

<p>This was false, India had the most, and as I learned it makes sense since their population is over 1 billion.</p>

<p>-America probably has the highest percentage as well.

</p>

<p>This was false as well and the most surprising one. Vatican City had the heist percentage despite the lowest population. </p>

<p>-Countries with fewer electronics (such as phones and computers) will have fewer users.

</p>

<p>This one is still a grey area that was left relatively unanswered. </p>

</body>

</html>