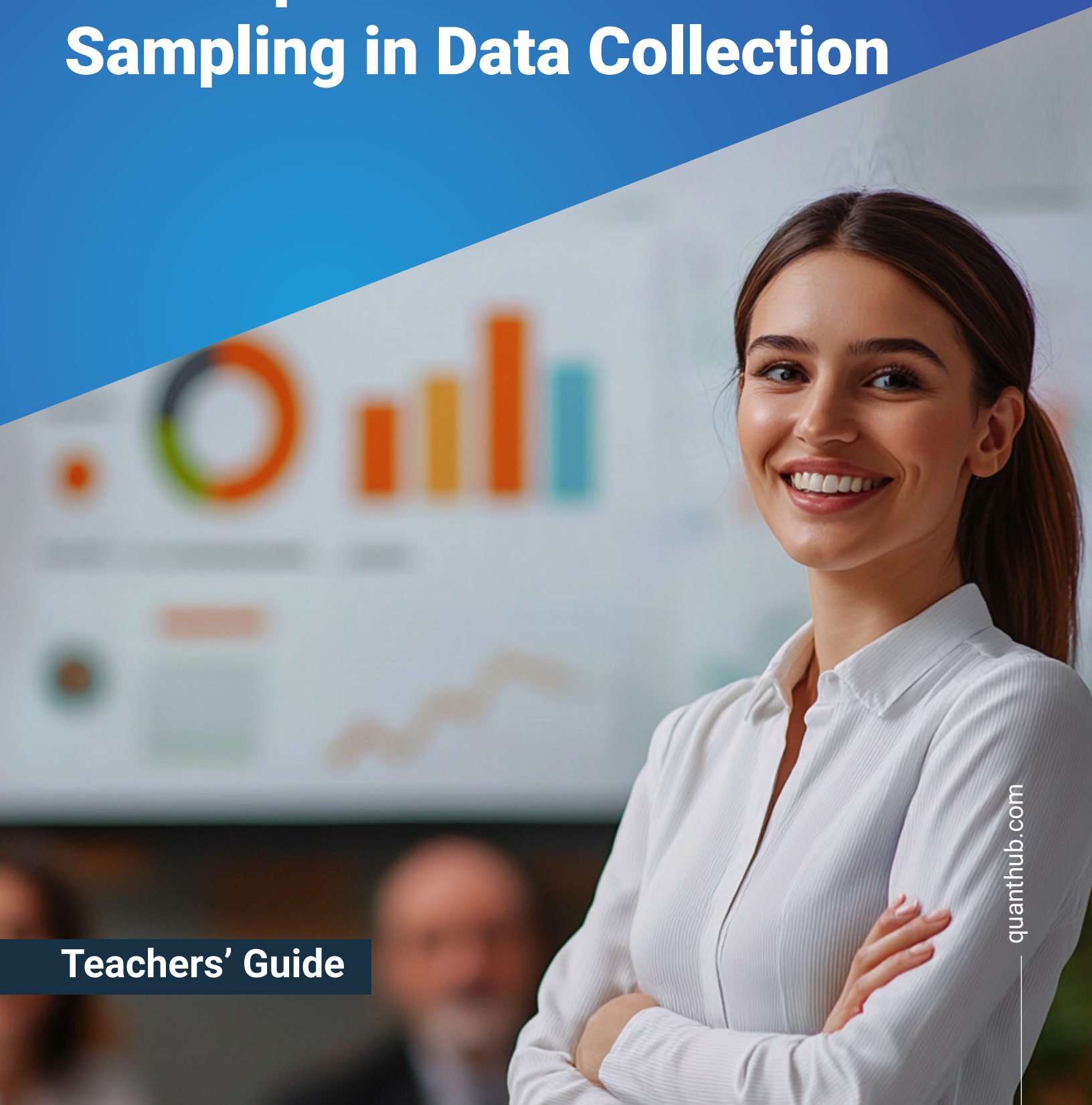
quanthub

The Importance of Random Sampling in Data Collection



Teachers' Guide

quanthub.com

Contents

Overview

The student will apply simple random sampling techniques to collect data, ensuring a fair and unbiased representation of the population of interest. The student will learn how to choose between sampling methods.

Standards

The student will apply the methods of sampling data, describing its importance in representing a broader population in a statistical study.

Learning Outcomes

The student will use simple random sampling to collect data

Duration

Approximately 60 minutes

Activity Titles	Total Time
INTRODUCE. Activity	10 minutes
EDUCATE. Activity	20 minutes
INSPIRE. Activity	30 minutes

Materials and Resources

Learning Resources:

General Requirements:

QuantHub Resources:

QuantHub Google Slides
QuantHub Dataset
QuantHub Article: The Power of Sampling

Required Background Knowledge

None

Introduce *(10 minutes)*

The activities in the Introduce phase are designed to help students make connections between past and present learning experiences, expose prior conceptions, and organize thinking toward the essential questions and learning outcomes of the learning sequence.

Strategy 1: Notice and wonder: Sampling candy

Instructions: Show a large jar filled with different colored candies. Ask the class to estimate how many of each color are in the jar. Then, take a small handful and count the colors. Discuss how this relates to sampling.

This hands-on activity introduces the concept of sampling in a fun, relatable way. It helps students understand how a small sample can represent a larger population.

Strategy 2: Think-pair-share: Everyday sampling

Instructions: Ask students to think of examples where they've seen or used sampling in their daily lives. Have them share with a partner, then discuss as a class.

This strategy activates prior knowledge and helps students connect the concept of sampling to their own experiences, making it more relevant and memorable.

Educate *(20 minutes)*

This activity helps Explain the topic to students in two parts. First, students are asked to share their initial models and explanations from experiences in the Introduce phase. Second, resources and information are provided to support student learning and introduce scientific or technological concepts. Students use these resources and information, as well as ideas of other students, to construct or revise their evidence-based models and explanations.

Strategy 1: Interactive lecture: Sampling scenarios

Instructions: Present various real-world scenarios and ask the class to decide if sampling is appropriate or not. Discuss the reasons behind each decision.

This strategy engages the class in critical thinking about when to use sampling, helping them understand the criteria for appropriate sampling situations.

Strategy 2: Collaborative learning: Sampling method matchup

Instructions: Divide the class into small groups. Give each group a set of cards describing different sampling methods and various research scenarios. Have them match the most appropriate sampling method to each scenario and explain their reasoning.

This hands-on activity helps students understand the different sampling methods and when to apply them, encouraging discussion and peer learning.

Strategy 3: Guided practice: Simple random sampling simulation

Instructions: Walk the class through a step-by-step process of conducting simple random sampling using a hypothetical research scenario. Use a spreadsheet or online tool to demonstrate each step.

This practical demonstration helps students understand the process of implementing simple random sampling, making the abstract concept more concrete.

Inspire *(30 minutes)*

These activities provide time for students to apply their understanding of concepts and skills. They might apply their understanding to similar phenomena or problems. This activity is meant to provide students an opportunity to reflect on the lesson and examine essential questions.

Scenario: You're a junior data analyst for the popular mobile game "Pixel Legends." The game developers want to improve player satisfaction and in-game purchases. They've asked you to gather data on player preferences, gaming habits, and spending patterns. With 100 active players, you need to use sampling techniques to collect this information efficiently.

Task: Determine if sampling is appropriate for this situation.
Task Instructions: Explain why sampling would be a good choice for gathering data about Pixel Legends players. Consider the population size, global distribution of players, and time constraints.

Task: Choose the best sampling method for surveying player preferences.
Task Instructions: Select the most appropriate sampling method (simple random, stratified, cluster, or systematic) for this scenario. Justify your choice based on the player base structure and the type of information you need to collect.

Task: Implement a simple random sampling technique to select players for the survey.
Task Instructions: Using the provided list of player ID numbers (1-100), select a sample of 35 players using simple random sampling. Describe each step of your process, including how you'll generate random numbers and select the players.

Assessment

Quick quiz or a worksheet to assess understanding of the empirical rule and normal distribution. Review and discuss the answers as a class.

Knowledge Check

Use the following QuantHub Scavenger Hunts to direct students' aRen9on to important terms, defini9ons, and ideas and quickly gauge students' understanding of concepts presented within this lesson.

Take the assessment:
https://upskill.quanthub.com/assessments/scavenger_hunts/skills/2185/start/1079

Review student progress:
https://upskill.quanthub.com/admin/assignment_reports/scavenger_hunts?skill_id=2185&learning_resource_id=1079

Additional Instructional Strategies

Acceleration

Activities designed to help advanced students or those who quickly grasp the material to move forward at a faster pace. The emphasis of these activities is on connecting students with additional external resources and applications of concepts to further their understanding real-world implications.

Intervention

Activities meant for students who are struggling with certain concepts or who need additional support to reach the expected level of understanding.

Appendix A: Introduce Script

Hello, everyone! Today, we're going to explore an exciting topic that's all around us, even if we don't always notice it. Let's start with a fun activity.

[Hold up the jar of candies]

Take a look at this jar. It's full of colorful candies. Now, I want you to guess: how many red candies do you think are in here? What about blue ones? Green?

[Allow a few students to share their guesses]

Great guesses! Now, here's the thing - it would take a long time to count every single candy, right? So, let's try something else.

[Take a small handful of candies from the jar]

I've just taken a small handful. Let's count these together.

[Count the candies by color]

We have 3 red, 2 blue, and 4 green. Now, do you think this small group tells us anything about what's in the whole jar?

[Allow for responses]

Exactly! This small group is what we call a 'sample'. It gives us an idea of what's in the whole jar without counting everything. This is called 'sampling', and it's super useful in many areas of life and science.

Now, let's think about where else we might see sampling. Turn to the person next to you and take a minute to think of times you've seen sampling in your everyday life. It could be anything!

[Allow 1-2 minutes for discussion]

Okay, let's hear some of your ideas. Who wants to share?

[Call on students to share. Possible responses might include:]

- Taste-testing a new recipe
- Polls on social media
- Product reviews online
- Weather forecasts

Those are all great examples! Sampling is everywhere. It helps us understand big groups by looking at smaller parts.

In our candy jar example, we used what's called 'simple random sampling'. We just grabbed a handful without trying to pick specific colors. This is one way to get a fair sample.

As we move forward, we'll learn more about how to use sampling to collect data. This is a key skill in many fields, from science to business to politics. It helps us make smart decisions based on information from large groups.

Remember, the next time you see a survey result or a product rating, you're looking at the result of sampling. It's a powerful tool that helps us understand our world better!

Any questions before we move on?

[Address any questions]

Great! Let's dive deeper into this fascinating topic.

Appendix B: Educate Script

Welcome, everyone! Today, we're diving deep into the world of sampling. Let's start with a fun activity to get our brains warmed up.

I'm going to describe some scenarios, and I want you to tell me if sampling would be a good idea or not. Ready? Here we go!

1. You're planning the school dance and want to know what music to play. Would sampling be a good way to find out?

[Allow for responses]

Great thoughts! This is a perfect scenario for sampling. We can't ask every single person in the school, but we can get a good idea by asking a smaller group. This is what we call a 'large population size' scenario.

2. The principal wants to know how many students have peanut allergies. Should we use sampling here?

[Allow for responses]

Interesting! This is actually a case where sampling might not be appropriate. We need to know about every single student for safety reasons. This falls under what we call 'complete population data required'.

Let's do a few more...

- A beverage company is planning to test a new energy drink flavor. They ask: "How can we gather meaningful feedback within our budget constraints?"

- A medical researcher is studying a rare genetic disorder with only 50 known patients. She wonders: "Is our patient pool large enough to justify sampling? How might sampling affect the validity of our results given the rarity of this condition?"
- A social media platform is considering a major redesign. They ponder: "How can we ensure feedback from all key user demographics?"

Excellent job, everyone! You're getting the hang of this. Now, let's talk about the different ways we can sample. We're going to play a little game called "Sampling Method Matchup".

[Divide the class into small groups]

Each group will get a set of cards. Some cards describe sampling methods, and others describe research scenarios. Your job is to match the best sampling method to each scenario. Don't worry if you're not sure - that's part of the learning process!

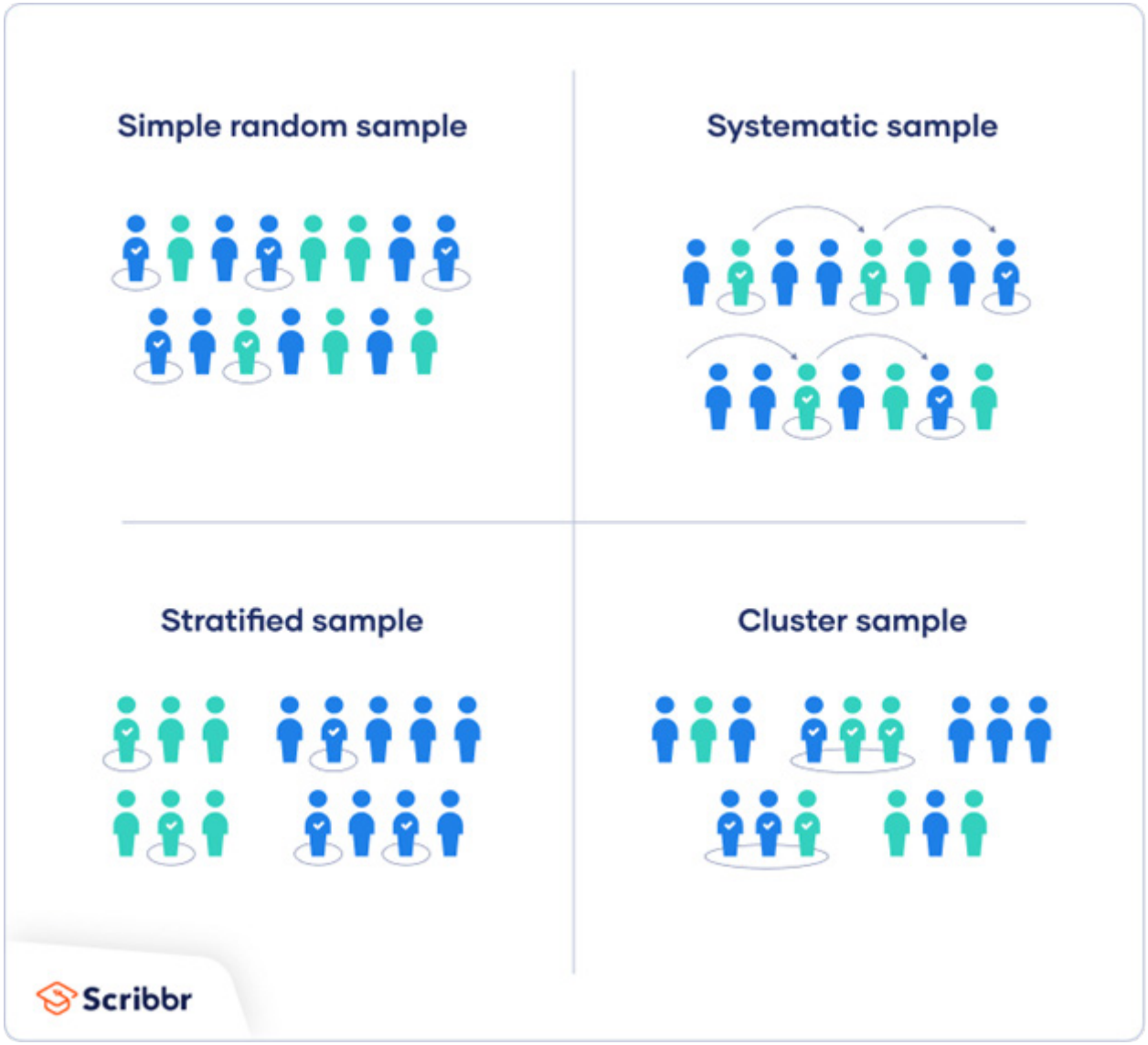
Cards:

- Stratified
- Simple
- Cluster
- Systematic

Scenarios:

- Planning school dance music selection
- Identifying students with peanut allergies
- Surveying all residents of a small town of 500 people
- Testing the effectiveness of a new drug
- Conducting a national census
- Studying voter preferences in a country with millions of voters
- Surveying employees in a large company with distinct departments
- Conducting a national health survey across a geographically large country
- Quality control inspection in a continuous production line

Here is a quick overview of each type:



Simple random sampling is like putting all the candies in a jar, shaking it well, and then picking out a few with your eyes closed. Every candy has an equal chance of being chosen. This method is fair and unbiased because you're not favoring any specific type of candy.

For example, let's say you have a list of 100 students, and you randomly select 10 names. Each student has an equal chance of being picked.

Stratified sampling is like organizing the candies by color first and then picking a few from each color group. This way, you ensure that every color is represented in your sample.

For example, the school has 500 students: 100 freshmen, 150 sophomores, 150 juniors, and 100 seniors. You pick 10 students from each grade. This ensures every grade level is represented in your sample.

Cluster sampling is like dividing the jar of candies into several smaller jars and then picking one whole jar to sample. Instead of picking individual candies from the whole jar, you randomly select one smaller jar to taste.

In the case of the school, your school has 20 classes, and you randomly select 3 entire classes to survey. Here, you're sampling all students within the chosen classes, not picking students individually.

Finally, systematic sampling is like lining up all the candies and picking every 10th candy. This method spreads your picks evenly throughout the jar.

So for the school, if you have a list of 100 students, and you select every 10th student on the list. If you start with the 5th student, you pick the 5th, 15th, 25th, and so on.

Each sampling method has its own advantages and is useful in different situations. Here's a quick recap:

- **Simple Random:** Best for fairness and unbiased results.
- **Stratified:** Ensures representation from all groups.
- **Cluster:** Convenient and cost-effective when groups are naturally formed.
- **Systematic:** Simple and quick to implement.

Okay, now that you have your overview, go ahead and get to work.

[Allow 10-15 minutes for groups to work]

Okay, let's hear what you've come up with. Who wants to share first?

[Discuss 2-3 matches, explaining the reasoning behind each]

Great work! You're really grasping these concepts. Now, let's put it all together and see how we actually do simple random sampling.

Imagine we're researchers studying study habits of high school students. We can't survey every high school student in the country, so we'll use simple random sampling. Here's how we do it:

1. First, we define our population. In this case, it's all high school students in the country.

2. Next, we need a sampling frame. Let's say we have a list of all 1,000 high school students.

3. Now, we need to decide our sample size. Let's say we want to survey 100 students.

4. Here's where the 'random' part comes in. We'll use a random number generator to select 100 numbers between 1 and 1000.

[Demonstrate using a spreadsheet or online tool]

5. The students corresponding to these numbers are our sample!

6. Finally, we collect data from these 100 students.

Sounds simple, right? But there are some challenges we need to watch out for. Can anyone think of potential problems?

[Allow for responses]

Great points! Some common issues are:

1. Ensuring true randomness - we need to use reliable random number generators.
2. Handling non-responses - what if some selected students don't participate?
3. Maintaining sample integrity - we need to stick to our plan and not make changes on the fly.
4. Balancing precision and practicality - we want accurate results, but we also have limited time and resources.

Remember, sampling is a powerful tool, but it's not always perfect. The key is understanding when and how to use it effectively.

Any questions before we wrap up?

[Address any questions]

Fantastic work today, everyone! You've taken a big step in understanding how researchers collect data in the real world. Next time you see a survey result or a study, you'll have a better idea of what's going on behind the scenes!

Appendix C: Inspire Activity Script

Assessment Instructions

1. Present the scenario and tasks to the class.
2. For individual assessment: Ask each person to complete the tasks independently, writing out their responses.
3. For group assessment: Divide the class into small teams (3-4 people). Each team should work together to complete the tasks, documenting their process and decisions.
4. Encourage showing all work, especially for the random sampling implementation.
5. After completion, lead a class discussion about the different approaches and solutions.
6. Use the provided rubric to assess understanding and application of sampling concepts.

Script

Hey, gamers! Today, we're diving into the world of Pixel Legends - not to play, but to level up our data skills. Imagine you're working for this hot mobile game, and they need your help to make it even better.

Your mission is to figure out how to gather info about what players like, how they play, and how they spend their money in the game. But here's the twist - you can't survey all million players.

That's where sampling comes in handy!

Let's break this down into three main quests:

First, determine if sampling is appropriate for this situation. Explain why sampling would be a good choice for gathering data about Pixel Legends players. Consider the population size, global distribution of players, and time constraints.

Next, choose the best sampling method for surveying player preferences. Select the most appropriate sampling method (simple random, stratified, cluster, or systematic) for this scenario. Justify your choice based on the player base structure and the type of information you need to collect.

Finally, implement a simple random sampling technique to select players for the survey. Using the provided list of player ID numbers (1-100), select a sample of 35 players using simple random sampling. Describe each step of your process, including how you'll generate random numbers and select the players.

Remember, sampling is like choosing the right characters for your team in a game. You can't use everyone, so you need to pick a group that represents all the players well.

As you work, think about why each choice matters. How might your decisions affect the info you get? What are the good and bad points of different approaches?

If you get stuck, don't rage quit! Think back to our class discussions and examples. And remember, in the real world, it's okay to ask for help or look things up. The important thing is understanding the process and making smart choices.

I'll be here if you need any hints or have questions. Good luck, and have fun with it! Who knows, maybe your survey will lead to the next big update in Pixel Legends!

Appendix D: Inspire Activity Excel and Python Instructions

Excel:

1. **Open your dataset:** Load your dataset into Excel.
2. **Add a helper column:** Insert a new column next to your data to generate random numbers. For example, if your data is in column A, insert a new column B.
3. **Generate random numbers:** In the first cell of the new column (let's say B2), enter the formula =RAND(). This function generates a random number between 0 and 1. Drag this formula down the column to fill it next to all your data rows.
4. **Sort by the random numbers:** Select your entire dataset, including the new column with random numbers. Then, go to the "Data" tab and click on "Sort." Sort your data by the column with the random numbers (e.g., column B) in ascending or descending order. This shuffles your dataset randomly.
5. **Select your sample:** After sorting, you can select the top N rows as your random sample. For example, if you want a random sample of 10 rows, select the first 10 rows of your shuffled dataset.

Example:

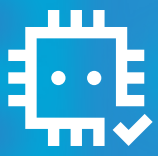
Assume your dataset is in column A (A1):

1. Insert a new column B.
2. In cell B2, enter =RAND().
3. Drag the formula down from B2 to B101.
4. Select columns A and B.
5. Go to "Data" -> "Sort" -> Sort by column B in ascending order.
6. Choose the top N rows from the sorted data as your random sample.

By following these steps, you can effectively create a random sample from your dataset in Excel.

Python:

```
import pandas as pd
import numpy as np

# Load your dataset
df = pd.read_csv('your_dataset.csv')  # Replace 'your_dataset.csv' with your file path

# Generate a random sample
# For a sample of 10 rows, you can use:
sample_df = df.sample(n=10, random_state=1)

# Display the sample
print(sample_df)

# Additional options:

# Fractional sampling (e.g., 10% of the dataset)
fraction_sample_df = df.sample(frac=0.1, random_state=1)
print(fraction_sample_df)
```



quanthub

k12.quanthub.com