

StatsSkills web app: a tool for systematically practising statistics skills

Damien Raftery
Dr Daire Ó Broin
Dr Denise Earle

StatsSkills Expression of Interest 



<https://forms.office.com/e/88SxWjNyCj>

Overview

- Demo
- Creating questions
- Initial pilot



Supported by the *National Forum for the Enhancement of Teaching and Learning* SATLE (Strategic Alignment of Teaching and Learning Enhancement) Initiative

StatsSkills Demo

Select *Skills* from *Skills Areas*

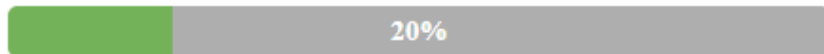
Home Skills Exercises Users (Admin) Skills (Admin) Exercises (Admin) Sign Out

SKILL AREAS

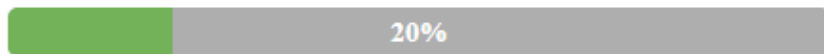
Descriptive Statistics Linear Correlation And Regression Percentages Probabilities Surveys And Samples Variable Types

First select the Skill Area to show the skills in the area. Then click on any skill that you would like to select. Click on the same skill again to unselect it. By selecting a skill, you will receive exercises tailored to that skill. The app will remember the skills you have previously selected – if you want to concentrate on different skills, remember to unselect the skills that you don't want to see questions on.

Correlation Coefficient



Prediction



Scatter



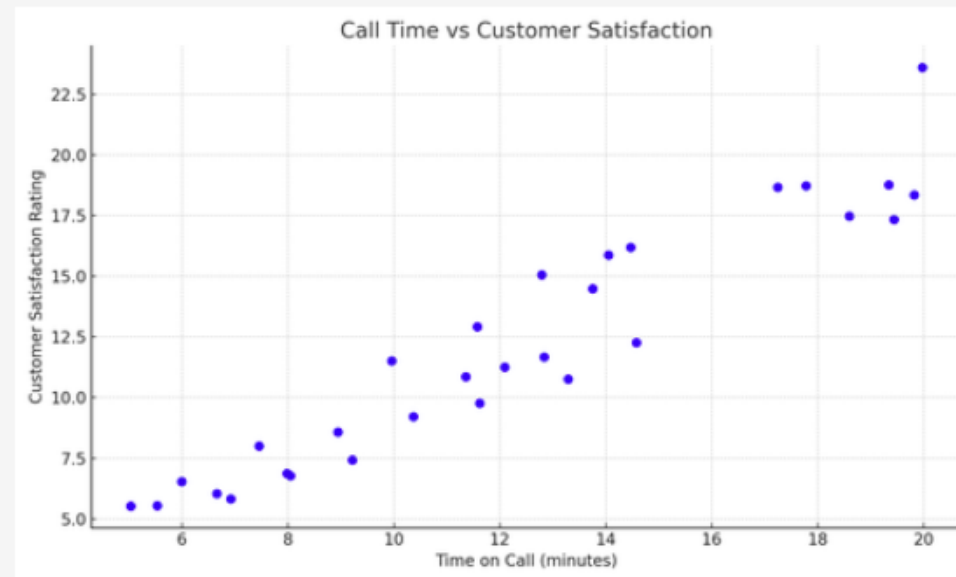
The Scatter Skill is selected

EXERCISES

Exercise 1/5

Random selection of 5 questions from selected Skills

A customer service department is reviewing data on the time spent on each customer service call and the subsequent customer satisfaction ratings. Which answer best describes the scatter diagram that would likely represent this data?



Your Answer:

- Perfect, negative linear correlation

Feedback for each question

Result:

Strong, positive linear correlation is indeed the correct answer!

Feedback:

The scatter diagram reveals that as website traffic increases, sales revenue also goes up, with data points closely aligned in an upward trajectory. This pattern suggests a strong positive correlation between website traffic and sales revenue.

Feedback customised to answer option chosen

Next Exercise

Prediction-type question

Exercise 1/5

A study on online advertising effectiveness shows that there is a strong positive linear relationship between the click-through rate (CTR) of online ads (x , in percentage) and the revenue generated from those ads (y , in thousands of euro). If the regression line is $y = 10 + 2x$, estimate the revenue generated when the ad's click-through rate is 8%. Enter your numerical answer in units of thousands of euro.

Your Answer:

16

Text input (or MCQ question type)

Submit

Result:

Unfortunately, '16' is incorrect... The correct answer is '26'.

Feedback:

Estimate the revenue generated by inserting the ad's click-through rate ($x=8$) into the regression line $y = 10 + 2x$, i.e. $y = 10 + 2(8) = 10 + 16 = 26$ thousands of euro.

StatsSkills Questions

610 Questions

Skill Areas →

Variable Types (129)

Descriptive Statistics (220)

Surveys & Samples (46)

Linear Regression & Correlation (102)

Percentages (50)

Probabilities (63)

Categorical or Numerical

Nominal, Ordinal, Continuous or Discrete

Mean, Median

Min, Max, Range

Quartiles, IQR

Describing Distributions

Sampling Methods

Question Types

Correlation Coefficient

Scatters

Prediction

One-way Tables

Two-way Tables

Simple

Intersection

Union

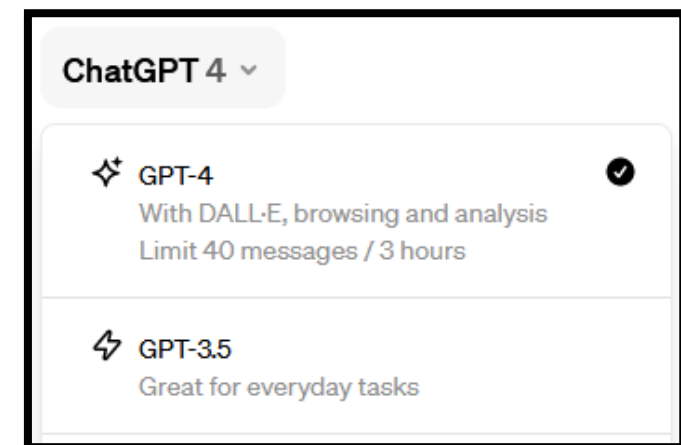
Conditional

Skills →



Question Design Process

- Current Skills & Skill Areas motivated by First Year Quantitative Techniques module in Bachelor of Business programme.
- Regular meetings between Denise & Damien (> 30 combined years of experience teaching introductory statistics/data analysis modules).
- Agreed on terminology, definitions & methods to use in question phrasing and feedback.
- Ensured questions were relatively quick for students to complete.
- Ensured questions covered all scenarios, e.g. when calculating the median, use odd and even numbers of digits, present digits sorted and unsorted.
- Used ChatGPT 4 to generate first draft of questions.



Using ChatGPT

- Asked ChatGPT to generate a small number of questions at first.
- Fine-tuned using additional prompts, e.g. expand on feedback, round answer to 1 decimal place.
- When questions were in suitable format, asked ChatGPT to generate:
 - Extra questions.
 - Excel file summarising the various details of the questions (uploaded to StatsSkills).
 - Links to download images.

You

Hi, I want to create several questions to assess first year college students' understanding of how to calculate the mean of a set of numbers. I want to have 5 or 6 numbers in each question. The numbers should be whole numbers (no fractions) and should have some context from the real-world, including questions that involve currency, specifically Euro. UK spelling and units of measurement should be used. After giving the question, I want to give the correct answer. I also want to give feedback that explains how to calculate the answer. Can you generate 10 such questions please?

ChatGPT

Of course! Here are 10 questions, along with their answers and feedback:

1. Question:

John scored the following marks in his 5 exams

Answer:

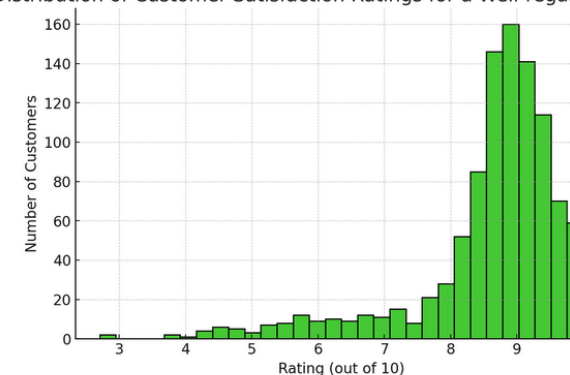
$$\text{Mean} = \frac{45 + 50 + 52 + 48 + 47}{5} = 48.4$$

Feedback:

To calculate the mean, add up all the numbers and divide by the number of numbers. In this case, the total is $45 + 50 + 52 + 48 + 47 = 242$. Dividing 242 by 5 gives 48.4.

ChatGPT

Distribution of Customer Satisfaction Ratings for a Well-regarded Product



Rating (out of 10)	Number of Customers
3	2
4	5
5	10
6	15
7	20
8	30
8	50
8	85
9	145
9	160
9	140
9	115
10	70
10	60

Here's the histogram illustrating the distribution of customer satisfaction ratings for a well-regarded product. As depicted, the majority of customers gave high ratings, but there's a pronounced tail on the left representing the fewer customers who weren't as satisfied, making the distribution clearly left-skewed.

You can [download the histogram here](#).

Using ChatGPT

Pros

- Saved us a lot of time!
- Generally provided an excellent variety of questions, numbers, real-world contexts, detailed feedback, etc...

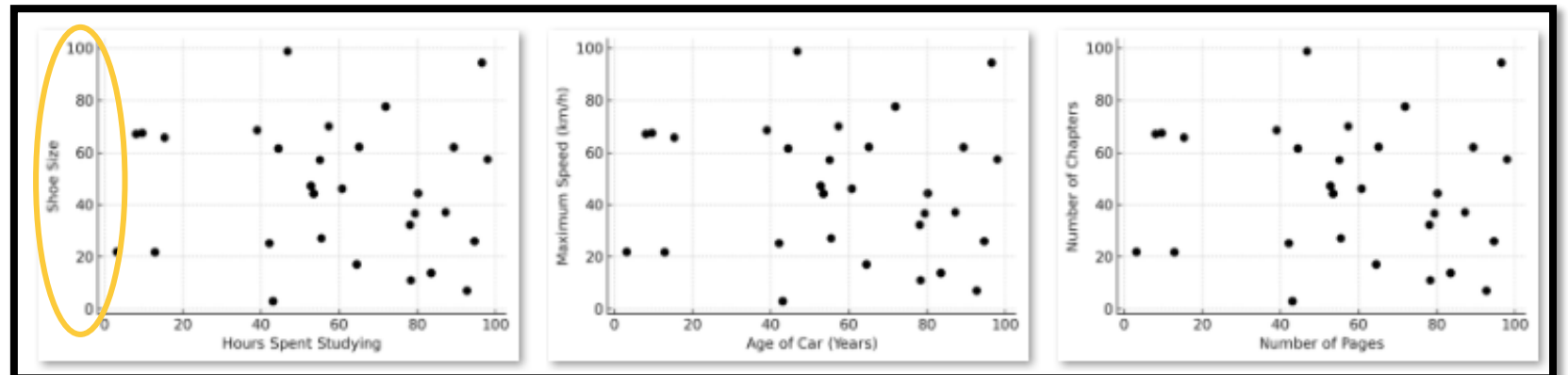
Question 2:

During a school's student council election, candidates were allowed to put up promotional posters. One candidate put up the following number of posters in nine different corridors: 20, 25, 30, 35, 40, 45, 50, 55, 60. How many posters did the candidate put up in the median corridor?

4. In a football tournament, five teams scored the following goals in a match: 87, 17, 82, 38, and 53. Which team scored the fewest goals?

Cons

- Generic feedback.
- Incorrect calculations (rare).
- Unrealistic numbers, e.g. Shoe Size = 100.
- Lazy:
 - Real-world contexts & digits became repetitive.
 - Identical scatterplots with different axis labels.
 - Obvious patterns in digits.
- **Results needed SIGNIFICANT proof-reading & lots of manual editing! Lecturers' expertise still required.**



StatsSkills pilot use (6 weeks)

81% used
(75 out of 93)

Questions answered

Correct	1,344	66%
Incorrect	678	34%

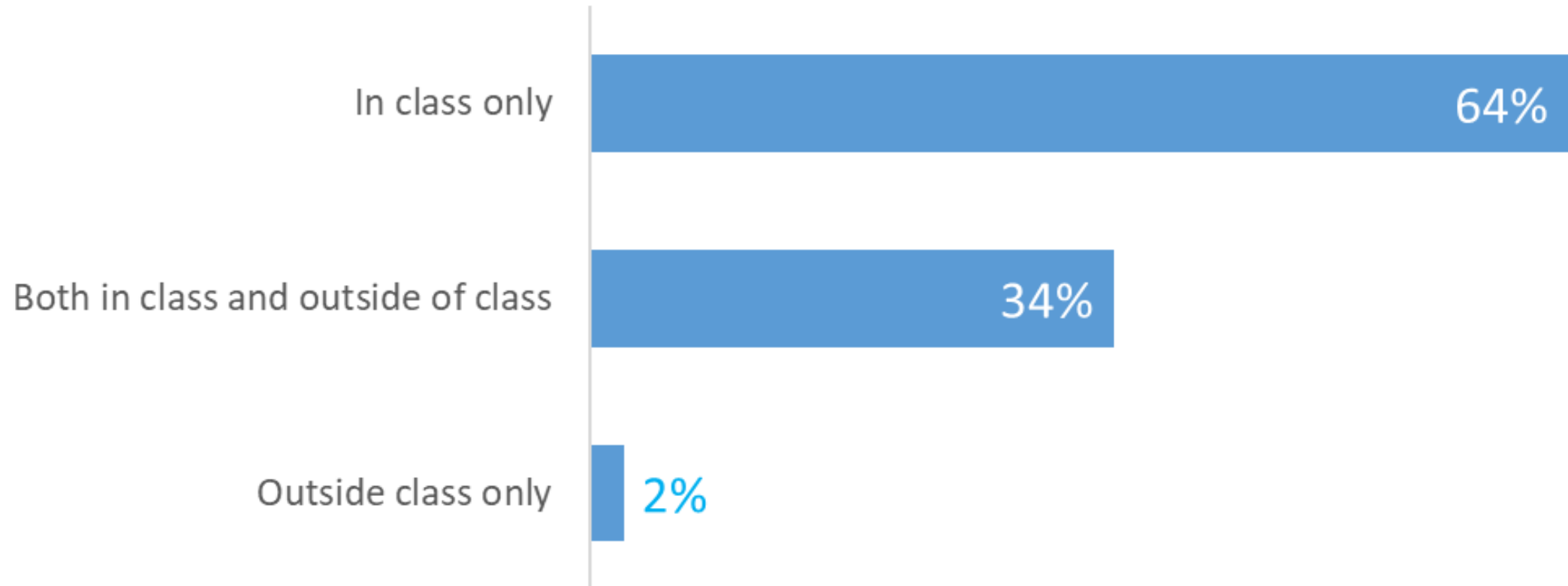
Total	2,022
No. of users	75
<hr/>	
Mean	27.0
Median	12
St dev	36.8
Min	1
Max	237

StatsSkills pilot survey

44 users (47%)

“Simple, handy and fun”

Setting(s) where used StatsSkills used



Number of times StatsSkills used



Strongly agree Agree Neutral Disagree Strongly disagree

I like using StatsSkills to learn

Using StatsSkills has increased my confidence on the topics covered

StatsSkills has a good mixture of questions on each skill

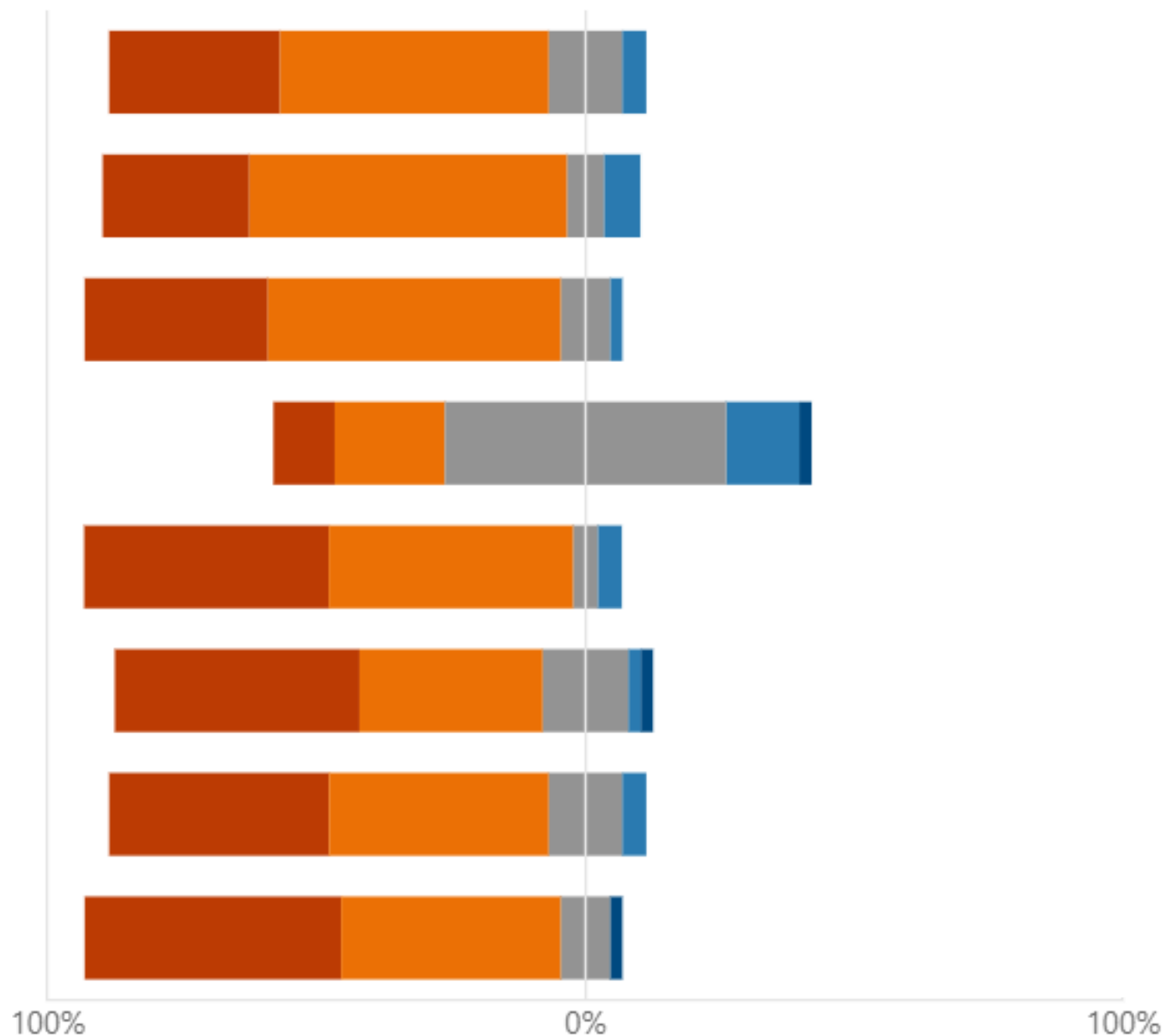
I learn more from studying my notes than using StatsSkills

The feedback I get after answering questions helps me to learn

I want my lecturers to add more skills areas

I want my lecturers to add more questions

StatsSkills is easy to use



Something you like about using StatsSkills?

straight forward topics exercise - easier Easy and clear Easily accessible
catagories Easy to navigate
small number question **easy** easy to use number of questions
Lots of question correct answer Easy to access Good for revision
practice questions right answer Simple and easy no timed

27 Comments:

Easy to use (14)

Feedback/explanations (8)

Helps practice (3)

“I appreciate the categories by topics,
as it simplifies the topics I need to focus on”

Something you suggest to improve StatsSkills?

issue of StatSkills
detailed questions option is wrong Downloadable app Wider range
questions in a quiz **questions** UI intuitive UI certain topic
range of questions Maybe adding no means
questions and topics Better lay Add more questions
categories reference to videos dealbreaker option is correct

17 Comments:

More topics/questions (9)

Better UI/layout (3)

StatsSkills web app: a tool for systematically practising statistics skills

Damien Raftery
Dr Daire Ó Broin
Dr Denise Earle

StatsSkills Expression of Interest 



<https://forms.office.com/e/88SxWjNyCj>