



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

200
1822 - 2022

Quantitative Research

Week 6

Man Geni Addysg Uwch yng Nghymru
Where Higher Education Began in Wales

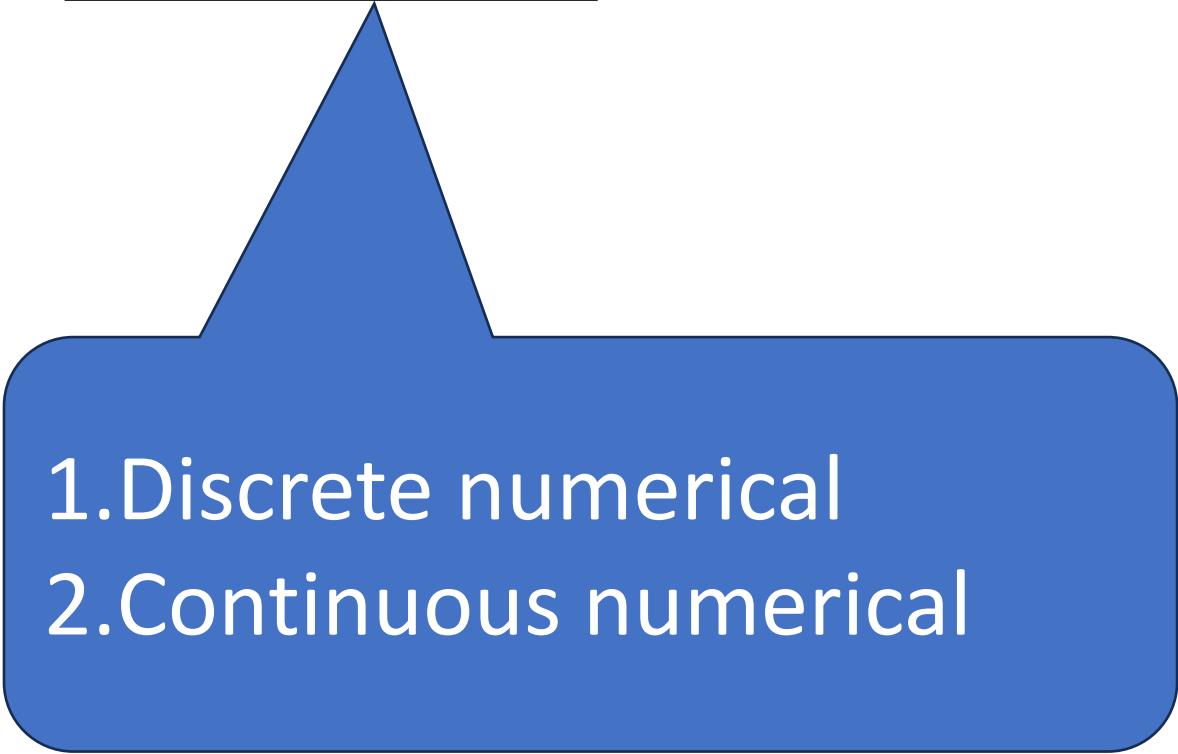
www.uwtsd.ac.uk

What is Quantitative Research?

- a method used to systematically collect and analyze **numerical data**
- to uncover patterns, test hypotheses, or measure relationships between variables.

Numerical data in Quantitative Research

- a method used to systematically collect and analyze numerical data

- 
1. Discrete numerical
 2. Continuous numerical

Numerical data in Quantitative Research

- a method used to systematically collect and analyze numerical data

1. Discrete numerical
2. Continuous numerical

Data that consists of
**countable, distinct
values**

Numerical data in Quantitative Research

- a method used to systematically collect and analyze numerical data

1. Discrete numerical
2. Continuous numerical

Data that consists of
**measurable values –
within a range**

Numerical data in Quantitative Research

An analyst is tracking the number of phishing emails received by a company each day.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Numerical data in Quantitative Research

An analyst is tracking the number of phishing emails received by a company each day.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Correct Answer: Discrete

The number of phishing emails is **countable** and consists of whole numbers (e.g., 5, 10, 15). You can't receive 3.5 emails.

Numerical data in Quantitative Research

Scenario:

A security operations center measures the time taken (in seconds) to respond to each security alert.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Numerical data in Quantitative Research

Scenario:

A security operations center measures the time taken (in seconds) to respond to each security alert.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Correct Answer: Continuous

Time is a **measurable** quantity that can take **decimal values** (e.g., 2.4 seconds), so it is continuous.

Numerical data in Quantitative Research

You are analyzing the number of unauthorized login attempts on a server over a month.

Question:

What type of data are you working with?

A. Continuous

B. Discrete

Numerical data in Quantitative Research

You are analyzing the number of unauthorized login attempts on a server over a month.

Question:

What type of data are you working with?

A. Continuous

B. Discrete

Correct Answer: Discrete

Login attempts are **counted** in whole numbers — you can't have 1.3 login attempts.

Numerical data in Quantitative Research

An IT admin logs the bandwidth usage (in Mbps) of a system throughout the day.

Question:

Which type of numerical data is this?

A. Continuous

B. Discrete

Numerical data in Quantitative Research

An IT admin logs the bandwidth usage (in Mbps) of a system throughout the day.

Question:

Which type of numerical data is this?

A. Continuous

B. Discrete

Correct Answer: Continuous

Bandwidth usage can have **decimal values** (e.g., 105.76 Mbps), and it is **measured**, not counted — making it continuous.

Numerical data in Quantitative Research

A cybersecurity researcher records the number of users who click on a malicious link in a simulated phishing test.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Numerical data in Quantitative Research

A cybersecurity researcher records the number of users who click on a malicious link in a simulated phishing test.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Correct Answer: Discrete

The number of users is a **countable** quantity and must be a **whole number** — you can't have 3.7 users.

Numerical data in Quantitative Research

A script records the time (in milliseconds) it takes for a server to respond to HTTP requests.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Numerical data in Quantitative Research

A script records the time (in milliseconds) it takes for a server to respond to HTTP requests.

Question:

What type of numerical data is this?

A. Continuous

B. Discrete

Correct Answer: Continuous

Milliseconds are part of a **continuous scale** — you can measure very small decimal differences in response time.

Numerical data in Quantitative Research

<u>Variable</u>	<u>Description</u>	<u>Data Type</u>
Number of malware infections	Count of infections per system or time period	Discrete numerical
Intrusion detection system (IDS) false positives	Frequency per day/week	Discrete numerical
Patch deployment time	Time (in hours or days) to apply patches	Continuous numerical
Encryption strength	Key length in bits (e.g., 128, 256)	Discrete numerical
Average login attempts before lockout	Per user or system	Continuous numerical
Network traffic volume	Megabytes or packets per second	Continuous numerical

Quantitative Research

to uncover patterns, **test hypotheses**, or measure relationships between variables.

IDS Effectiveness

- H_0 (Null Hypothesis): There is no difference in detection rates between IDS A and IDS B.
- H_1 (Alternative Hypothesis): IDS A has a significantly higher detection rate than IDS B.

Test: Use a two-sample t-test comparing detection rates (as percentages) across a sample of networks.

Quantitative Research

to uncover patterns, test hypotheses, or measure relationships between variables.

IDS Effectiveness

- H_0 (Null Hypothesis): There is no difference in detection rates between IDS A and IDS B.
- H_1 (Alternative Hypothesis): IDS A has a significantly higher detection rate than IDS B.

Test: Use a two-sample t-test comparing detection rates (as percentages) across a sample of networks.

Quantitative Research

to uncover patterns, test hypotheses, or measure relationships between variables.

IDS Effectiveness

- H_0 (Null Hypothesis): There is no difference in detection rates between IDS A and IDS B.
- H_1 (Alternative Hypothesis): IDS A has a significantly higher detection rate than IDS B.

Test: Use a two-sample t-test comparing detection rates (as percentages) across a sample of networks.

Quantitative Research

to uncover patterns, test hypotheses, or measure relationships between variables.

IDS Effectiveness

- H_0 (Null Hypothesis): There is no difference in detection rates between IDS A and IDS B.
- H_1 (Alternative Hypothesis): IDS A and B have different detection rates.

Test: ?

Quantitative Research

to uncover patterns, test hypotheses, or measure relationships between variables.

IDS Effectiveness

- H_0 (Null Hypothesis): There is no difference in detection rates between IDS A and IDS B.
- H_1 (Alternative Hypothesis): IDS A and B have different detection rates.

Test: two-tailed test, requiring different critical values)

Quantitative Research

Key points

- H_0 = 'no effect' baseline; H_1 = research claim
- Choose tests based on data type and hypothesis structure
- Always validate test assumptions.

Quantitative Research

to uncover patterns, test hypotheses, or **measure relationships between variables.**

Correlation

Regression

Measures the *strength* and *direction* of the relationship between two variables.

Quantitative Research

to uncover patterns, test hypotheses, or **measure relationships between variables.**

Correlation

Variables:

Hours Studied (X)

Exam Score (Y)

Quantitative Research
to uncover patterns, test hypotheses, or measure relationships between variables.

Variables:

Hours Studied (X)

Exam Score (Y)

Correlation:

If Hours Studied  and Exam Score  → Positive Correlation (e.g., +0.85).

If Hours Studied  but Score  → Negative Correlation (e.g., -0.60).

If no pattern → Near Zero Correlation (e.g., +0.10).

Quantitative Research

to uncover patterns, test hypotheses, or **measure relationships between variables.**

Correlation

Regression

Predicts the value of one variable (Y) based on another (X).

Quantitative Research

to uncover patterns, test hypotheses, or **measure relationships between variables.**

Regression

Variables:

House Size (X) in sq. ft.

House Price (Y) in \$

$$\text{Price} = \$50,000 + \$300 \times \text{Size}$$

$$\text{If Size} = 1,000 \text{ sq. ft.} \rightarrow \text{Predicted Price} = \$50,000 + \$300 \times 1000 = \$350,000$$

Quantitative Research

to uncover patterns, test hypotheses, or **measure relationships between variables.**

Regression

Variables:

House Size (X) in sq. ft.

House Price (Y) in \$

$$\text{Price} = \$50,000 + \$300 \times \text{Size}$$



$$Y = a + bX$$

Quantitative Research

Regression

Research Question: *Does increasing firewall rule complexity (number of rules) impact network latency?*

• **Variables:**

- Independent: Number of firewall rules
- Dependent: Network latency (ms)

Method:

Linear regression

Output: $\text{Latency} = \beta_0 + \beta_1 \times (\text{Number of rules}) + \varepsilon$

Quantitative Research

Correlation

Research Question: Is there a relationship between number of phishing emails received and number of reported incidents?

Data: 30 organizations, each reporting number of phishing emails and incidents per month

Test: Pearson correlation

Result: A correlation coefficient $r = 0.78$ suggests a strong positive relationship

Quantitative Research

Correlation

Research Question: Is there a relationship between number of phishing emails received and number of reported incidents?

Data: 30 organizations, each reporting number of phishing emails and incidents per month

Test: Pearson correlation

Result: A correlation coefficient $r = 0.78$ suggests a strong positive relationship

Quantitative Research

One-way ANOVA

A statistical test used to **compare the means of 3 or more groups** to determine if at least one group is significantly different from the others.

Quantitative Research

One-way ANOVA

- **Research Question:** *Do organizations of different sizes (small, medium, large) experience different average numbers of DDoS attacks per year?*
- **Variables:**
 - Independent: Organization size (categorical)
 - Dependent: Number of DDoS attacks (numerical)

Quantitative Research

One-way ANOVA

- **Research Question:** *Do organizations of different sizes (small, medium, large) experience different average numbers of DDoS attacks per year?*
- **Variables:**
 - Independent: Organization size (categorical)
 - Dependent: Number of DDoS attacks (numerical)

Organization Size	DDoS Attacks (per year)
Small	2, 3, 1, 2, 3
Medium	5, 6, 7, 5, 6
Large	10, 12, 9, 11, 13

Group	Mean	Std. Dev
Small	2.2	0.84
Medium	5.8	0.84
Large	11.0	1.58

Quantitative Research

Source	SS	df	MS	F	p-value
Between Groups	216.13	2	108.07	91.99	<0.001
Within Groups	17.60	12	1.47		
Total	233.73	14			

p-value	Interpretation
< 0.05	Statistically significant (reject H_0)
≥ 0.05	Not statistically significant (fail to reject H_0)

Quantitative Research

Key takeaway

Use correlation to:

- Screen variables for further analysis.
- Check if two trends move together (e.g., GDP growth and employment).

Use regression to:

- Forecast sales based on ad spending.
- Quantify how diet affects cholesterol levels.

Activity: Quantitative Research in Cybersecurity

Research Question: Do organizations of different sizes experience different average numbers of DDoS attacks per year?

Activity: Quantitative Research in Cybersecurity

Null hypothesis (H_0): There is no difference in the average number of DDoS attacks across organization sizes.

Alternative hypothesis (H_1): There is a difference in the average number of DDoS attacks across at least one group.

Activity: Quantitative Research in Cybersecurity

Input the Dataset:

Organization Size	DDoS Attacks (per year)
Small	2
Small	3
Small	1
Small	2
Small	3
Medium	5
Medium	6
Medium	7
Medium	5
Medium	6
Large	10
Large	12
Large	9
Large	11
Large	13

Activity: Quantitative Research in Cybersecurity

Calculate Group Statistics

In Excel, for each group (**Small, Medium, Large**), calculate:

- Mean
- Standard Deviation
- Count (n)

You can use formulas like:

=AVERAGE(range)

=STDEV.P(range)

=COUNT(range)

Activity: Quantitative Research in Cybersecurity

Perform One-Way ANOVA

In Excel:

1. Go to **Data** → **Data Analysis** → **ANOVA: Single Factor**
2. Input the **range of data** grouped by columns or rows
3. Select **Output Range** and click OK

Activity: Quantitative Research in Cybersecurity

Perform One-Way ANOVA

Interpret Results

Record the ANOVA output below:

Term	Value
F-statistic	_____
p-value	_____

Activity: Quantitative Research in Cybersecurity

Make a Conclusion

Is the **p-value** < 0.05 ?

Yes $\rightarrow$ Reject the null hypothesis

No $\rightarrow$ Fail to reject the null hypothesis

Write your conclusion:

What does this mean for cybersecurity teams in small vs large organizations?

ANOVA

Excel interface showing an ANOVA analysis of phishing attacks across organization sizes.

Organization ID

Organization ID	Organization Type	Number of Phishing Attacks (Last Year)	Summary Stats
1	Small	12	Organization Type
2	Small	15	Mean
3	Small	10	SD
4	Medium	20	Count
5	Medium	18	
6	Medium	22	
7	Large	33	
8	Large	30	
9	Large	28	

Anova: Single Factor

Groups	Count	Sum	Average	Variance
Small	3	17.84994	5.949982	30.6188
Medium	3	25	8.333333	102.3333
Large	3	35.84994	11.94998	253.5191

ANOVA

Source of SS	df	MS	F	P-value	F crit
Between Groups	2	27.38026	0.21254	0.814363	5.143253
Within Groups	6	128.8238			
Total	8				

"A one-way ANOVA showed no statistically significant difference in phishing attack frequency across organization sizes, $F(2, 8) = [Insert F]$, $p = 0.814$. Therefore, we fail to reject the null hypothesis."

XLMiner Analysis ToolPak

Anova: Single Factor

Input Range:

Grouped By: ☐ Columns ☒ Rows

☒ Labels in First Column

Alpha:

Output Range:

Anova: Two-Factor With Replication

Anova: Two-Factor Without Replication

Correlation

Covariance

Descriptive Statistics

Exponential Smoothing

F-Test Two-Sample for Variances

Fourier Analysis

Histogram

Linear Regression

Logistic Regression

Moving Average

Random Number Generation

Anova

Output range missing. Please enter

Quantitative Research – ANOVA

AutoSave Off Book1 - Excel

File Home Insert Draw Page Layout Formulas Data Review View Automate Help

Clipboard Font Alignment Number Styles Cells Editing Sensitivity Add-ins Commands Group

XLMiner Analysis ToolPak

- Anova: Single Factor
- Anova: Two-Factor With Replication
- Anova: Two-Factor Without Replication
- Correlation
- Covariance
- Descriptive Statistics
- Exponential Smoothing
- F-Test Two-Sample for Variances
- Fourier Analysis
- Histogram
- Linear Regression
- Logistic Regression
- Moving Average
- Random Number Generation
- Rank and Percentile
- Sampling
- t-Test: Paired Two Sample for Means
- t-Test: Two-Sample Assuming Equal Variances
- t-Test: Two-Sample Assuming Unequal Variances
- z-Test: Two-Sample for Means

Copyright © 2025 Frontline Systems, Inc.

<https://forms.office.com/e/eHEsgh5nbr>



Coming up...

Qualitative Method