

RESEARCH METHODS

Lecture Title: Making Sense of Numbers
(Quantitative Analysis)

Lecturer: Ms Mary Anne Puli

Date: 13th February 2026



**Masters in
Human Resources Management**

What Today's Lecture Is About

- We will learn how to *understand and interpret* numerical data.
- The aim is to reduce anxiety around statistics and build your confidence.
- You do **not** need to be a statistician to follow this lecture.
- We focus on plain English explanations that support your assignments and dissertation.

Why this matters

- Masters-level research requires you to analyse data, not only describe it.
- Good data interpretation helps you make stronger, evidence-based conclusions in HRM.



Tools for Quantitative Analysis

Software you will use

- **Excel** – basic functions for simple analysis.
- **SPSS** – the most common tool in HRM and organisational behaviour research (OBR).
- Often free to download for a limited period.
- **Other tools exist**
- Matlab, Stata, and R are also strong statistical packages.

In this lecture, we will focus on **SPSS** because it is widely used in HRM research.



Quantitative Data Analysis

Don't panic about statistics

- Many HR students feel overwhelmed by technical terms.
- But statistics are simply about *describing things with numbers* and *assessing the odds* (Rugg & Petre, 2007).
- You do **not** need to understand every statistical method.
- Use only what is relevant to your research.

Our approach today

- We will explain key terms in **plain English**.
- Focus on understanding, not memorising formulas.
- Build confidence to interpret data and make informed HR decisions.
- **Remember:**

You are not expected to become a statistician.

You are learning how to use data meaningfully in HRM.



Introducing SPSS

Masters in
Human Resources Management

What Are Variables?

Variables

- A variable is anything that can change or vary.
- In HR research, variables can relate to:
 - Individuals' (e.g. age, commitment, performance)
 - Teams (e.g. team performance)
 - Organisations (e.g. size, sector)

Why this matters

- You will use variables in every piece of quantitative research.
- Understanding them helps you build clear research questions and hypotheses.



Independent and Dependent Variables

Independent variable (IV)

- The “cause” or influencing factor.
- Example: High levels of absence from work.

Dependent variable (DV)

- The “effect” or outcome that changes because of the independent variable.
- Example: Low productivity levels.

Directionality

- We assume the IV comes first (antecedent) and the DV follows (outcome).
- This is a theoretical assumption unless tested over time.



Independent and Dependent Variables

Masters in

Human Resources Managment



Cross-Sectional vs Longitudinal Data

Longitudinal design

- Data collected at two or more points in time.
- Allows you to test whether the cause truly comes before the effect.
- Very difficult to achieve in student research projects.

Cross-sectional design (most common)

- Data collected at one moment in time.
- You cannot claim causality.
- But you *can* discuss directionality based on theory.
- Example: High absence levels are **negatively related** to productivity.



What Is a Hypothesis?

Hypothesis

- An assumption or an *informed guess* about the relationship between two or more variables.
- It guides your analysis and helps you test whether a relationship exists.
- Example: *High levels of absence from work are negatively related to productivity levels.*

Why it matters

- Quantitative analysis is built around testing hypotheses.
- Hypotheses help you make clear, focused claims in your research.



The Null Hypothesis

Null hypothesis (H_0)

- States that there is **no relationship** between the variables.
- Example: *There is no relationship between absence levels and productivity.*

Why we use it

- Significance tests compare your data against the null hypothesis.
- If the data allow you to reject the null hypothesis, you can say a significant relationship exists.

Key idea

- You are not proving your hypothesis is “true”.
- You are showing that the null hypothesis is unlikely to be correct.



The Null Hypothesis explained

Null Hypothesis

DEFINITION

EXAMPLES

STRENGTHS

WEAKNESSES

NULL HYPOTHESIS
VS ALTERNATIVE
HYPOTHESIS

The null hypothesis is a fundamental concept in statistics and scientific research.



What is Probability in Quantitative Analysis?

Probability (p-value)

- Helps us judge whether our findings happened by *chance* or reflect a real pattern.
- A p-value is a number between **0 and 1**.
- The closer the p-value is to **0**, the more confident we can be that the result is meaningful.
- In HR research, a p-value **below 0.05** is usually considered strong evidence *against* the null hypothesis.

Why this matters

- Probability helps you decide whether a relationship between variables is significant.
- It supports evidence-based HR decisions.



Understanding Common p-Values

Key thresholds

$p = 0.05 \rightarrow$ about **1 chance in 20** that the result happened by chance.

$p = 0.01 \rightarrow$ about **1 chance in 100** that the result happened by chance.

$p = 0.001 \rightarrow$ about **1 chance in 1,000** that the result happened by chance.

Interpretation

- Lower p-values mean stronger confidence in your findings.
- If the p-value is below 0.05, you can usually reject the null hypothesis and say the relationship is significant.



What Do We Mean by “Significance”?

Significance

- A way of expressing how confident we are that our findings are *not* due to chance.
- It tells us whether results from our sample are likely to apply to the wider population.
- Significance is based on **probability (p-values)**.

Key idea

- Lower p-values mean stronger confidence in your findings.
- Significance helps you discuss your results clearly and professionally in HR research.



Levels of Significance

Accepted expressions

- **Significant**
 - About *1 chance in 20* that the result happened by chance
 - $p \leq 0.05$
- **Highly significant**
 - About *1 chance in 100* that the result happened by chance
 - $p \leq 0.01$
- **Very highly significant**
 - About *1 chance in 1,000* that the result happened by chance
 - $p \leq 0.001$
- **How to use this in your writing**
 - “The relationship between absence and productivity was **significant** ($p < 0.05$).”
 - “Differences in engagement scores were **highly significant** ($p < 0.01$).”

One-Tailed and Two-Tailed Hypotheses

One-tailed hypothesis

- You predict a relationship in a **specific direction**.
- Example:
 - *Stress* → *absence from work*
 - You expect stress to **increase** absence.
- The arrow has **one point**, so it is *one-tailed*.

Two-tailed hypothesis

- You expect a relationship, but the **direction is unclear**.
- Example:
 - *Manager self-confidence* ↔ *manager training*
 - Training might increase confidence, or confident managers might seek more training.
- The arrow has **two points**, so it is *two-tailed*.



Why Tails Matter in Statistical Testing

- Statistical software (e.g. SPSS) calculates p-values differently depending on whether your hypothesis is one-tailed or two-tailed.
- Most software reports **two-tailed p-values** by default.

Interpreting p-values

- If you have a **one-tailed** hypothesis, you normally divide the two-tailed p-value by **two**.
- Two-tailed tests are more common in HR research because direction is often uncertain.
- **Key point**
- One-tailed = directional prediction
- Two-tailed = relationship could go either way





What Are Parametric and Non-Parametric Data?

Parametric data

- Data that meets certain criteria.
- If your data is parametric, you can use specific statistical tests (e.g. t-tests, ANOVA, regression).

Non-parametric data

- Data that does *not* meet those criteria.
- You simply use different tests (e.g. Mann–Whitney, Kruskal–Wallis, Spearman's correlation).
- Non-parametric tests are not “worse” — they just work differently.
- **Key message**
- This is only a way of categorising data.
- It should not cause stress — both types of tests are valid and widely used.



How Do You Know If Your Data Is Parametric?

Your data is **parametric** if you can answer **YES** to all of these:

- Is your data **normally distributed**?
- Are your **group sizes similar** if you are comparing groups?
- Did you ask **all participants the same questions**?
- Are your measures on an **interval or ratio scale**?
- If you answer **NO** to one or more:
- Use **non-parametric** tests instead.
- This is completely fine — non-parametric tests still give strong, meaningful results.



Examples of Interval and Ratio Scales in Surveys

Interval scale (no true zero, equal intervals)

- Measures feelings, attitudes or perceptions on a scale.
- **Survey example:**
 - *“How satisfied are you with your job?”*
 - 1 = Very dissatisfied
 - 2 = Dissatisfied
 - 3 = Neutral
 - 4 = Satisfied
 - 5 = Very satisfied

You can say the difference between 2 and 3 is the same as between 4 and 5, but you cannot say “5 is twice as satisfied as 2”.

Ratio scale (true zero, equal intervals, can multiply/divide)

- Measures quantities with a meaningful zero point.
- **Survey examples:**
 - *“How many years have you worked in this organisation?”*
 - *“How many hours of training did you complete last month?”*
 - *“How many days were you absent in the past year?”*
- Here, 0 means “none”, and 10 years is twice as much as 5 years.



Parametric vs Non-Parametric Data

Practical Tips for Managing Your Data Analysis

Masters in
Human Resources Management

Stay organised

- Back up your work frequently — every 20 minutes is sensible during analysis.
- Keep a clear record of:
 - Which tests you have run
 - Why you ran them
 - What still needs to be done
- Label SPSS outputs logically so you can find them again.

Check your work carefully

- Double-check numbers, outputs and scoring.
- If a result looks “too good to be true”, it probably is.
- Make sure all scale items run in the **same direction** before analysing.

Choose the right tests

- Match the test to your **data type** (nominal, ordinal, interval, ratio).
- Use **non-parametric tests** if your data is not normally distributed.
- Larger samples make significance easier to achieve — interpret p-values cautiously.

Present results clearly

- Copy SPSS tables directly into Word to avoid typing errors.
- Use charts sparingly — only when they genuinely add value.
- Quality matters more than quantity.



An Analysis Route map – Where Do I Start?

Step 1: Review your research questions

- What do you want to understand?
- What problem are you trying to solve?
- Which variables matter?

Step 2: Review your data

- Look at frequencies and descriptive statistics.
- Spot possible patterns or associations.
- Identify surprises or things you need to check.
- Start forming hypotheses.

Step 3: Choose the right statistical test

- Are you comparing groups?
- Are you testing relationships between scale variables?
- Are you exploring whether items cluster into factors?

Your choice of test depends on the type of variables and the purpose of your analysis.



Choosing the Right Test

If your data is parametric

- Use **parametric tests**, which are more powerful when assumptions are met:
 - **t-tests** → compare two groups
 - **ANOVA** → compare three or more groups
 - **Pearson correlation** → test associations
 - **Regression** → test prediction

If your data is non-parametric

- Use **non-parametric tests**, which do not assume normal distribution:
 - **Mann–Whitney U** → compare two independent groups
 - **Wilcoxon signed-rank** → compare paired groups
 - **Spearman's rho** → test associations
 - **Non-parametric regression** → prediction without normality

Summary

- Start with your **research questions**.
- Check your **data type** and **distribution**.
- Then select the **appropriate test**.

Descriptive Statistics using SPSS

Masters in
Human Resources Management

Inferential Statistics Explained

Masters in
Human Resources Managment



Assessing the relationship between two variables

Bivariate analysis

- This means examining the **relationship between two variables**.
- It helps you see whether the variables move together, differ between groups or show no association.
- If there is **no relationship**, the variables are described as **independent**.

Why start here?

- It is one of the simplest forms of quantitative analysis.
- Ideal for students who feel nervous about statistics.
- Helps you build confidence before moving to more complex tests.

Examples of bivariate questions

- Do men and women differ in engagement levels?
- Is perceived organisational support related to employee engagement?
- Does tenure predict job satisfaction?



Probability and Significance

What Is Cross-Tabulation?

- **Cross-tabulation (contingency tables)**
- A method for examining how **two categorical variables** relate to each other.
- Shows how cases in one category are distributed across categories of another variable.
- Very useful for **nominal data** (e.g. gender, age group, qualification level).
- Helps you spot patterns, differences or independence between groups.

Why use it?

- Easy to interpret visually.
- Helps you see whether groups differ in meaningful ways.
- As we already saw it is a good starting point before running more advanced tests.

Reading a Cross-Tabulation Table

What the table shows

- Counts (how many people fall into each cell).
- Percentages (what proportion of each group falls into each category).
- Percentages make comparison easier, especially with large samples.
- **How to interpret it**
- Look for **direction**: do percentages increase or decrease across groups?
- If percentages look **roughly equal**, there may be **no relationship**.
- If one group has noticeably higher or lower percentages, there may be an **association**.
- **Important note**
- If a group is **small (less than 50)**, raw numbers may be more reliable than percentages.



Creating and Testing Cross-Tabs in SPSS

How to build a cross-tab in SPSS

- Analyse → Descriptive Statistics → Crosstabs
- Place one variable on **Rows** and the other on **Columns**
- Select **Cells** → **Percentages** if you want SPSS to calculate them
- Click **OK**

Next step: Testing significance

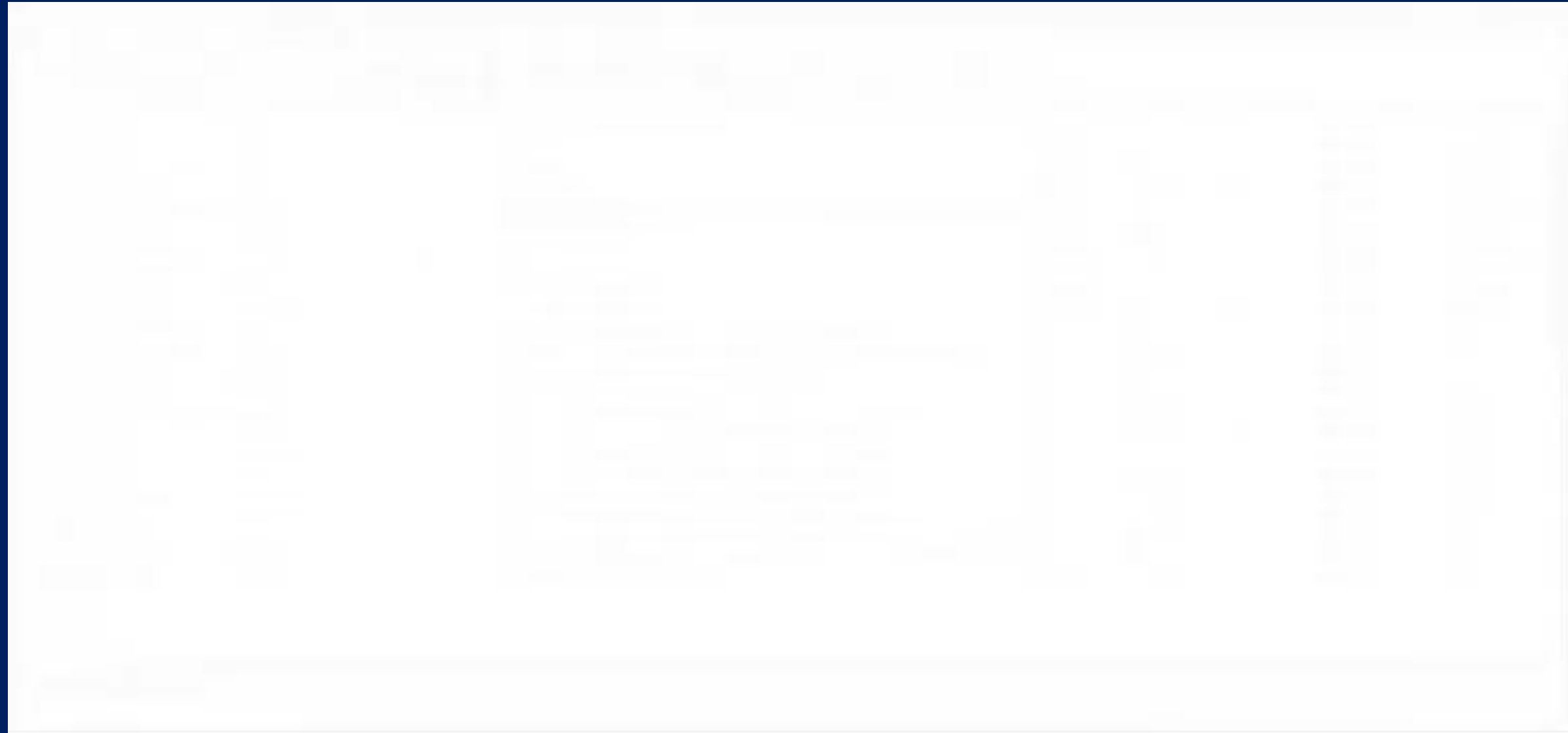
- To check whether the relationship is statistically meaningful, run a **Chi-Squared test**.
- Chi-Squared helps you test the **null hypothesis** that the variables are independent.
- If the Chi-Squared result is significant, the variables are **not independent**.

Why this matters

- Cross-tabs show the pattern.
- Chi-Squared tells you whether the pattern is **unlikely to be due to chance**.



Using Percentages in Cross-Tabulations



Reading the table

- Look at the **percentages** within each group.
- Ask: Do the percentages move in a clear direction?
 - If yes → there may be a **relationship**.
 - If no → the variables may be **independent**.

Why percentages help

- With large samples (e.g. 1,500+ respondents), percentages make patterns easier to see.
- They show the **proportion** of each group in each category.
- **Testing significance**
- To check whether the pattern is **real** and not due to chance, run a **Chi-Squared test**.
- Chi-Squared helps you test the **null hypothesis** that the variables are independent.
- A **significant** result means the variables are **related**.

What is the Chi-Squared Test? (χ^2)

Purpose of the test

- Used with **cross-tabulations** (tables comparing two categorical variables).
- Helps you check whether the two variables are **independent** or **related**.
- In simple terms: **Is the pattern in the table real, or could it have happened by chance?**

How it works

- Compares what you **expected** to see (if there was no relationship) with what you **actually** observed in the table.
- If the difference is large, the variables are **not independent**.



Why Probability and Significance Matter

Cross-tabulations

- Great for spotting patterns between two categorical variables.
- But they **cannot** tell you whether the pattern might have happened **by chance**.
- **Why we test significance**
- We need to know whether the association we see is **real** or just coincidence.
- This is where **probability (p-values)** and **significance tests** come in.

The Chi-Squared test (χ^2)

- The most common test for cross-tabulated data.
- Used widely in business and management research.
- Helps you check whether the rows and columns in your table are **independent** or **related**.



How the Chi-Squared Test Works

What the test does

- Calculates what the counts *should* be in each cell if there was **no relationship** (expected values).
- Compares these with the **actual counts** in your table.
- The bigger the difference, the more likely the variables are **not independent**.

Interpreting the result

- If the Chi-Squared test is **significant ($p < 0.05$)**:
 - You can reject the **null hypothesis** (no relationship).
 - The variables are likely **associated**.
- If it is **not significant**:
 - The variables are probably **independent**.

Important note

- Chi-Squared is only reliable when **expected counts are at least 5** in each cell.
- If you find a significant association, you may then explore the **strength** of the relationship using a **correlation test**.



Running a Chi-Squared Test in SPSS

How to run the test

- Analyse → Descriptive Statistics → Crosstabs
- Put one variable on **Rows**
- Put the other variable on **Columns**
- Click **Statistics** → tick **Chi-square**
- Press **Continue** → **OK**

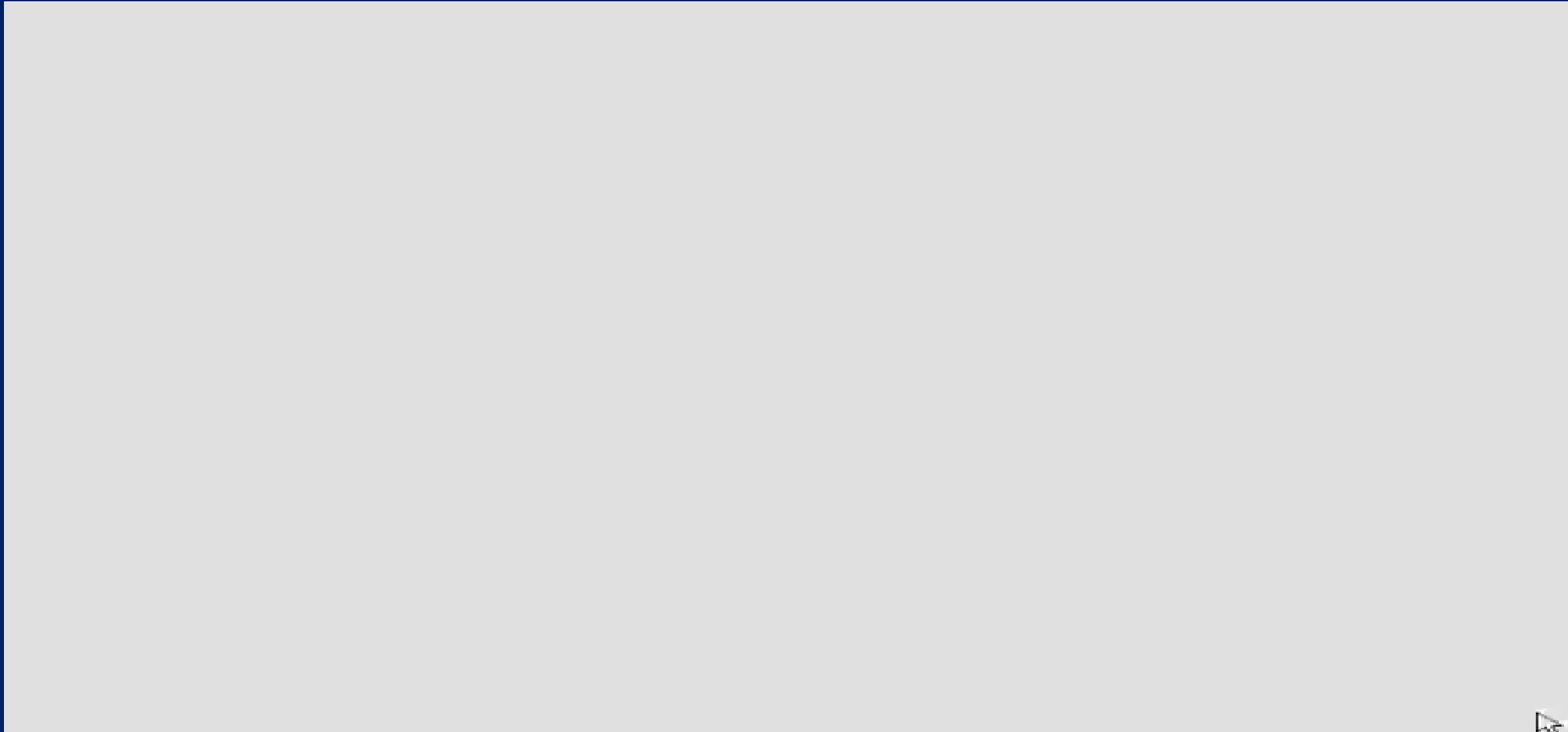
What SPSS will give you

- A cross-tabulation table
- A Chi-Squared test table showing:
 - **Pearson Chi-Square value**
 - **Degrees of freedom (df)**
 - **Asymp. Sig. (2-sided)** → this is your **p-value**

Why we use it

- To check whether the relationship between two categorical variables is **real** or just **due to chance**.

Chi-Squared using SPSS



What Are Tests of Correlation?

Purpose of correlation tests

- To measure the **strength** and **direction** of the relationship between two variables.
- They show whether variables move **together**, **oppositely** or **not at all**.
- Useful when you want to see whether two things vary in a **systematic** way.
- **Types of relationships**
- **Positive correlation**
 - Both variables increase or decrease together
 - Example: More classroom training ↔ more coaching investment
- **Negative correlation**
 - One variable increases while the other decreases
 - Example: Higher stress ↔ lower job satisfaction
- **No correlation**
 - The size of one variable does not relate to the other
- **Important reminder**
- Correlation shows **association**, not **causation**.
- It cannot tell you whether one variable *causes* the other.



Pearson, Spearman and Interpreting Correlation Strength

- Two main correlation tests
- Pearson's r
 - Used for **parametric** data
 - Measures linear relationships
- Spearman's ρ (r_s)
 - Used for **non-parametric** data
 - Measures ranked or monotonic relationships
- Correlation coefficient values
- Range from **-1 to +1**
 - **+1** = perfect positive relationship
 - **0** = no relationship
 - **-1** = perfect negative relationship
- How to interpret strength (when $p = 0.000$)
- **0.2 to 0.4** → a *good* level of association
- **0.4 to 0.6** → a *very strong* association in business and management research
- Key point
- A significant correlation (e.g., $p = 0.000$) means the relationship is **unlikely to be due to chance**, but it still does **not** prove causation.



Correlation tests using SPSS

Masters in
Human Resources Management

What is Regression Analysis?

Purpose of regression

- Regression helps you test whether **one variable predicts another**.
- The variable doing the predicting is the **independent variable**.
- The variable being predicted is the **dependent variable**.
- Useful when you want to understand **influences, drivers or antecedents** in HR research.

Types of data needed

- Works with **interval, ordinal or ratio** scale data.
- Nominal data can be used **only if it is dichotomous** (e.g. manager vs non-manager; male vs female).

Types of regression

- **Simple regression**: one predictor → one outcome
- **Multiple regression**: several predictors tested **at the same time**
 - Common in HRM research because many factors influence behaviour
- **Key point**
- Regression shows **prediction**, not **proof of causation**.



When Can You Use Multiple Regression?

Sample size requirements

- Regression needs a **large sample** to be reliable.
- Your sample must be much bigger than the number of predictors.
- Minimum guideline:
 - **5:1 ratio** → at least 5 participants per predictor
- Better practice:
 - **10:1 ratio** → 10 participants per predictor
- Some researchers prefer:
 - **40:1 ratio** for stronger reliability

Why sample size matters

- Too few cases can produce unstable or misleading results.
- A larger sample gives more accurate estimates of how strongly each predictor relates to the outcome.

When not to use regression

- If your sample is **small**, regression is **not appropriate**.
- Choose simpler tests (e.g. correlation or group comparisons) instead.



How to Run a Multiple Regression in SPSS

- Analyse → Regression → Linear
- Put your **dependent variable** (outcome) into the *Dependent* box
- Put your **independent variables** (predictors) into the *Independent(s)* box
- Choose a method:
 - **Enter** (adds all predictors at once)
 - **Stepwise** (adds predictors in steps based on statistical criteria)
- Click **Statistics** and tick:
 - *Model fit*
 - *Descriptives*
 - *Estimates*
 - *R squared change* (only if not using Enter)
- Press **Continue** → **OK**
- **What SPSS will produce**
- Six tables, including:
 - Descriptive statistics
 - Correlations
 - Model summary (R^2)
 - ANOVA
 - Coefficients (β values and significance)

Tip

- Don't be overwhelmed by the output — you only need to focus on a few key tables.



Why Use Multiple Regression?

Why not just test one predictor?

- Example hypothesis: *Higher perceived organisational support → higher employee engagement.*
- But engagement may also be influenced by:
 - Gender
 - Age
 - Education
 - Tenure
 - Managerial role
- These are **control variables** — alternative explanations you need to rule out.

How to include control variables

- Use the **Stepwise** method:
 - **Step 1:** Add control variables (gender, age, education, tenure, managerial role)
 - **Step 2:** Add your main predictor (perceived organisational support)
 - Dependent variable = employee engagement

Why this matters

- It shows whether your main predictor still explains engagement **after** accounting for other factors.
- This strengthens your argument and makes your findings more credible.

Important reminder

- Regression needs a **large sample**.
- Ideally at least **10 participants per predictor** (more is better).



Linear Regression by SPSS

Masters in
Human Resources Management

What Are Tests of Difference?

Purpose of tests of difference

- Used to check whether **two groups** give **different responses** to the same question.
- Helpful when you want to compare groups such as:
 - People from different countries
 - Managers vs non-managers
 - Employees at two different time points

Which test should you use?

- **Parametric data → t-test**
- **Non-parametric data → Mann–Whitney test**
 - Often used in HR Masters projects because data is rarely perfectly parametric.

Types of t-tests

- **Independent samples t-test**
 - Groups are unrelated (e.g. employees from two countries).
- **Paired samples t-test**
 - Same people measured twice (e.g. before and after training).
 - Sample size must match in both measurements.

Non-parametric equivalent

- **Wilcoxon signed-rank test**
 - Used for paired groups when data is non-parametric.



Choosing the Right Test

Independent groups

- Two groups with **no connection** between members.
 - Example: comparing attitudes of employees in Country A vs Country B
- Use:
 - **Independent samples t-test** (parametric)
 - **Mann–Whitney U test** (non-parametric)

Paired groups

- The **same people** measured twice, or naturally linked pairs.
 - Example: employee survey at Time 1 and Time 2
 - Example: 35 couples → 35 men paired with 35 women
- Use:
 - **Paired samples t-test** (parametric)
 - **Wilcoxon signed-rank test** (non-parametric)
- **Key reminders**
 - Always check whether your data is **parametric or non-parametric**.
 - For paired tests, the number of cases must be **identical** in both groups.
 - These tests help you decide whether group differences are **real** or just **due to chance**.



Running an Independent Samples t-Test in SPSS

Masters in
Human Resources Management

When to use this test

- You want to compare **two independent groups**
 - Example: Managers vs Non-managers
 - Example: Country A vs Country B
- Your data is **parametric** (normally distributed, interval/ratio scale)

Steps in SPSS

- Analyse → Compare Means → Independent Samples T Test
- Put your **test variable(s)** (e.g. engagement scores) into the *Test Variable(s)* box
- Put your **grouping variable** (e.g. managerial role) into the *Grouping Variable* box
- Click **Define Groups**
 - Enter the two values used in your dataset (e.g. 1 = Managerial, 0 = Non-managerial)
- Click **Continue** → **OK**

What SPSS will produce

- Group statistics (means and standard deviations)
- The t-test table showing:
 - t-value
 - Degrees of freedom (df)
 - Significance (p-value)



Interpreting the Independent Samples t-Test

What to look for

- **Group means**
 - Shows which group scored higher or lower
- **p-value (Sig. 2-tailed)**
 - Tells you whether the difference is statistically significant

How to interpret the p-value

- **$p < 0.05$**
 - The difference between the two groups is **significant**
 - The groups responded differently
- **$p \geq 0.05$**
 - No significant difference
 - Any difference in means may be due to chance

Example

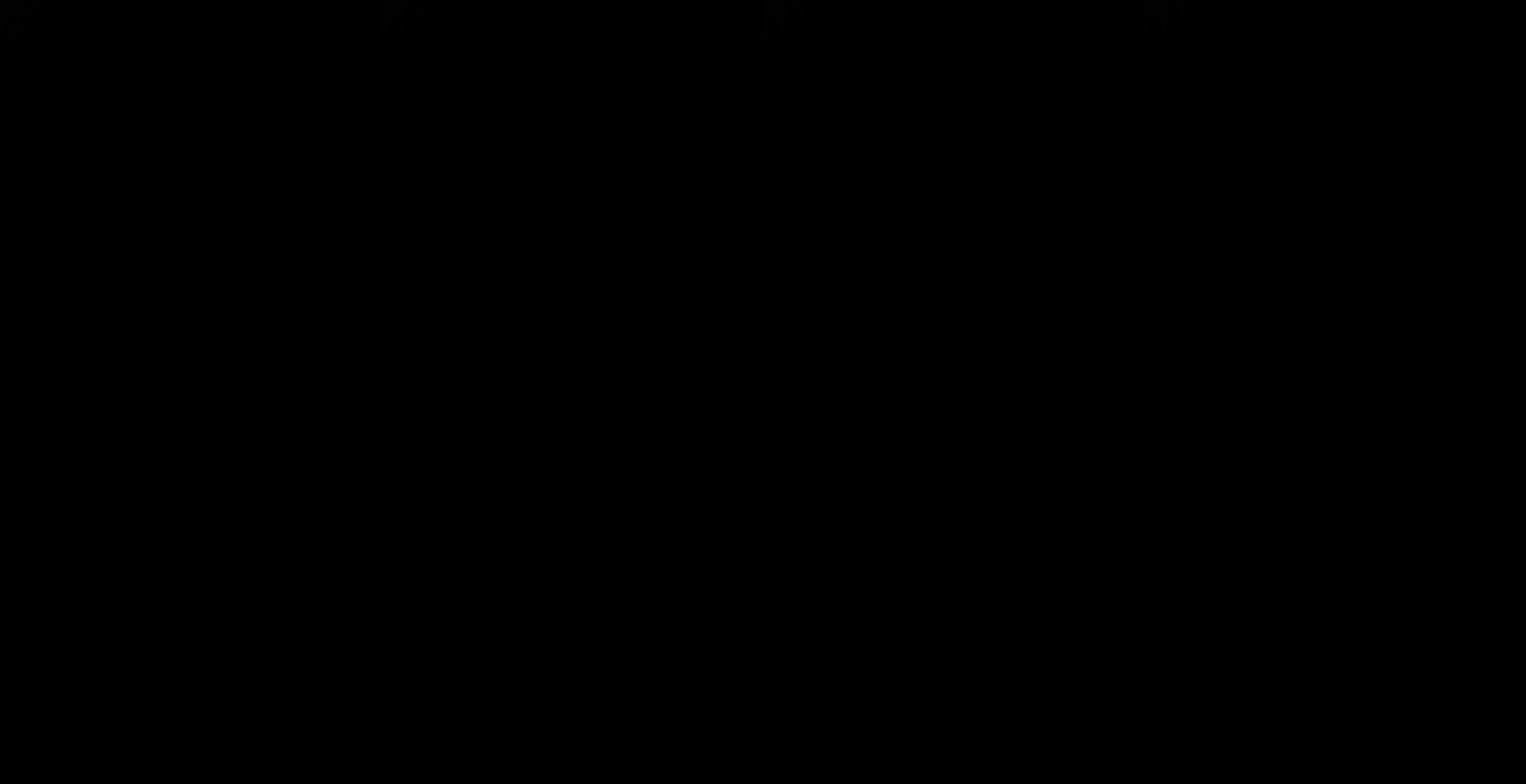
- If managers score higher on engagement than non-managers
- And the p-value is **0.03**
 - You can say: *"Managers report significantly higher engagement than non-managers."*

Key reminder

- This test only works for **independent groups**.
- If the same people are measured twice, you must use a **paired samples t-test** instead.

One Sample t-Test

Masters in
Human Resources Management



What Is Factor Analysis and Why Do We Use It?

Why HR researchers need it

- Many HR concepts (e.g. empowerment, engagement, leadership style) are **latent constructs**.
- These cannot be observed directly, so we measure them using **multiple questions**.
- But not all questions are equally good indicators of the construct.

What factor analysis does

- Helps you see whether your questions form **patterns** or **clusters**.
- These clusters represent meaningful sub-dimensions of a construct.
- Example: Engagement → vigour, dedication, absorption.

Two main uses

- To check whether your items group together into **coherent constructs**.
- To see whether your items suggest the presence of **key underlying variables**.

Summary: Factor analysis helps you check whether your survey questions actually measure what you think they measure.



How Factor Analysis Helps

The challenge

- When you have many questions, how do you know which ones truly measure your construct?
- Some items may be strong indicators; others may be weak or irrelevant.

What factor analysis does

- A statistical technique that groups items based on how strongly they relate to each other.
- Helps you identify **clusters** (factors) that represent meaningful sub-dimensions.
- Shows which items belong together and which do not.

Why it is useful

- Helps you build **clean, reliable scales**.
- Ensures you are measuring the construct you intended to measure.
- Commonly used in HR research when working with surveys like UWES, POS, leadership scales, etc.

Summary

- Latent constructs cannot be measured directly.
- You need multiple questions.
- Factor analysis helps you check whether those questions actually work together to measure the same thing.



How Factor Analysis Works

What the technique looks for

- It examines how strongly each question is **correlated** with every other question.
- From these correlations, it identifies:
 - **Clusters** of items that belong together
 - **How much variation** in the data is explained by each cluster (factor)

What this tells you

- Whether your items form clear, meaningful groups
- Whether some items are weak indicators and should be removed
- Whether your scale is **valid** (measures the right thing) and **reliable**

Related techniques

- **Factor Analysis** and **Principal Component Analysis (PCA)**
 - Different mathematical approaches
 - But both help you identify underlying structures in your data

Why this matters in HR research

- Ensures your measurement tools are **sound**
- Strengthens the quality of your findings
- Helps you build better, more accurate scales for complex constructs



How to Run a Factor Analysis in SPSS

Steps in SPSS

- Analyse → Data Reduction → Factor
- In the *Extraction* box:
 - Choose **Principal Components** (a common starting point)
 - Tick **Scree plot** to help decide how many factors to keep
- In the *Rotation* box:
 - Select **Varimax** (helps make factor patterns clearer)
- Click **Continue** → **OK**

What SPSS will produce

- A long output with several tables and graphs
- Key elements include:
 - Correlation matrix
 - Total variance explained
 - Scree plot
 - Rotated component matrix (shows which items load onto which factor)

Tip

- Don't worry about the complex maths behind factor analysis.
- Focus on the **patterns**: which items cluster together, and how much variance each factor explains.



Interpreting Factor Analysis Output

What the output tells you

- Whether your survey items form **clear clusters** (factors)
- How strongly each item relates to each factor (factor loadings)
- How much of the overall variation in responses each factor explains

Variance explained

- SPSS shows the **percentage of variance** accounted for by each factor
- Different disciplines have different expectations
- In HRM research, because human behaviour is complex:
 - A factor explaining **over 40% of the variance** is considered **interesting and meaningful**

Why this matters

- Helps you check whether your items genuinely measure the construct you intended
- Supports the **validity** of your scale
- Helps you decide which items to keep, revise, or remove

Summary: Factor analysis helps you see whether your questions “hang together” in a sensible way and whether they truly measure the underlying concept you care about.

Factor Analysis and Principal Components Analysis in SPSS

Masters in
Human Resources Management

Satisfaction With Life Scale_Factor Analysis Example_updated.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

9: SWLS4 5.00 Visible: 5 of 5 Variables

	SWLS1	SWLS2	SWLS3	SWLS4	SWLS5	var	var	var	var	var	var	var	var
1	6.00	6.00	6.00	6.00	5.00								
2	5.00	6.00	6.00	6.00	6.00								
3	5.00	6.00	6.00	4.00	5.00								
4	4.00	3.00	3.00	4.00	3.00								
5	2.00	6.00	6.00	6.00	6.00								
6	6.00	7.00	5.00	5.00	3.00								
7	6.00	5.00	6.00	6.00	6.00								
8	6.00	3.00	5.00	5.00	6.00								
9	5.00	4.00	4.00	5.00	6.00								
10	6.00	6.00	6.00	6.00	6.00								
11	5.00	7.00	7.00	7.00	7.00								
12	5.00	7.00	6.00	2.00	6.00								
13	7.00	7.00	7.00	7.00	7.00								
14	5.00	6.00	6.00	5.00	5.00								
15	3.00	4.00	5.00	5.00	3.00								
16	4.00	4.00	6.00	5.00	6.00								
17	6.00	6.00	6.00	6.00	6.00								
18	6.00	5.00	7.00	6.00	6.00								
19	3.00	4.00	4.00	5.00	3.00								
20	1.00	1.00	1.00	1.00	2.00								
21	5.00	6.00	6.00	3.00	5.00								
22	5.00	5.00	5.00	5.00	4.00								

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:OFF

Principal Components Analysis in SPSS

*Satisfaction With Life Scale_Factor Analysis Example_updated.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Visible: 5 of 5 Variables

	SWLS1	SWLS2	SWLS3	SWLS4	SWLS5	var	var	var	var	var	var	var	var
1	6.00	6.00	6.00	6.00	5.00								
2	5.00	6.00	6.00	6.00	6.00								
3	5.00	6.00	6.00	4.00	5.00								
4	4.00	3.00	3.00	4.00	3.00								
5	2.00	6.00	6.00	6.00	6.00								
6	6.00	7.00	5.00	5.00	3.00								
7	6.00	5.00	6.00	6.00	6.00								
8	6.00	3.00	5.00	5.00	6.00								
9	5.00	4.00	4.00	5.00	6.00								
10	6.00	6.00	6.00	6.00	6.00								
11	5.00	7.00	7.00	7.00	7.00								
12	5.00	7.00	6.00	2.00	6.00								
13	7.00	7.00	7.00	7.00	7.00								
14	5.00	6.00	6.00	5.00	5.00								
15	3.00	4.00	5.00	5.00	3.00								
16	4.00	4.00	6.00	5.00	6.00								
17	6.00	6.00	6.00	6.00	6.00								
18	6.00	5.00	7.00	6.00	6.00								
19	3.00	4.00	4.00	5.00	3.00								
20	1.00	1.00	1.00	1.00	2.00								
21	5.00	6.00	6.00	3.00	5.00								
22	5.00	5.00	5.00	5.00	4.00								

Data View Variable View

Variables IBM SPSS Statistics Processor is ready Unicode:OFF

Thank
you
for
your
attention!

