



# Getting Started with IBM SPSS Statistics

**Jarlath Quinn – Analytics Consultant**

[www.sv-europe.com](http://www.sv-europe.com)

A SELECT INTERNATIONAL COMPANY



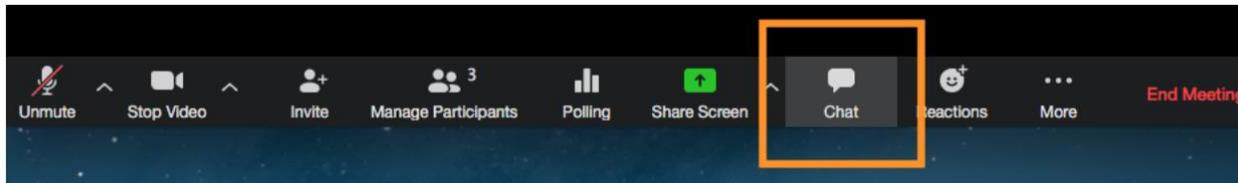
# Getting Started with IBM SPSS Statistics

**Jarlath Quinn – Analytics Consultant**

**Just waiting for all attendees to join...**

# FAQ's

- Is this session being recorded? Yes
- Can I get a copy of the slides? Yes, we'll email links to download materials after the session has ended.
- Can we arrange a re-run for colleagues? Yes, just ask us.
- How can I ask questions? All lines are muted so please use the chat panel – if we run out of time we will follow up with you.





- Gold accredited partner to IBM, Predictive Solutions and DataRobot specialising in advanced analytics & big data technologies
- Work with open-source technologies (R, Python, Spark etc.)
- Team each has 15 to 30 years of experience working in the advanced and predictive analytics industry
- Deep experience of applied advanced analytics applications across sectors
  - Retail
  - Gaming
  - Utilities
  - Insurance
  - Telecommunications
  - Media
  - FMCG



# Agenda

- An overview of SPSS Statistics
- Reading data from databases and spreadsheets
- Formatting the data
- Creating filters and banded fields
- Using SPSS Tables functionality to create publication-quality reports
- Finding statistically significant differences in groups

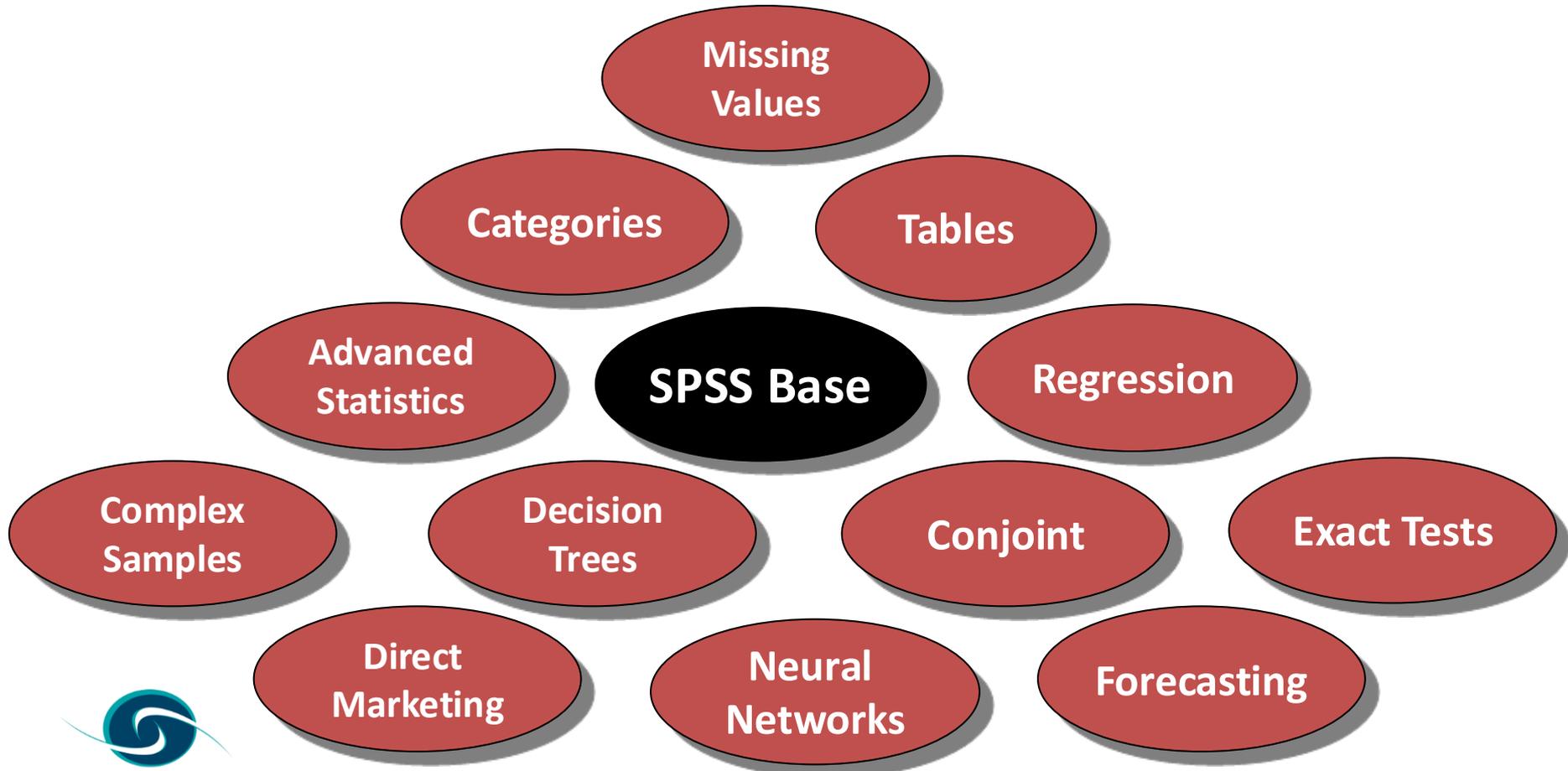


# An Overview of IBM SPSS Statistics

# SPSS Statistics

- Since 1968 one of the world's most popular data analysis and statistical interrogation platforms
- Used for everything from epidemiology studies , survey research and business reporting to direct marketing, credit risk, process modelling and asset management
- Statistics included in the base software:
  - Descriptive statistics: Cross tabulation, Frequencies, Descriptives statistics
  - Statistical Tests: T-test, ANOVA, Correlation
  - Prediction for numerical outcomes: Linear regression
  - Identifying relationships and groups: Factor analysis, Cluster analysis

# IBM SPSS Statistics Base & Associated Modules

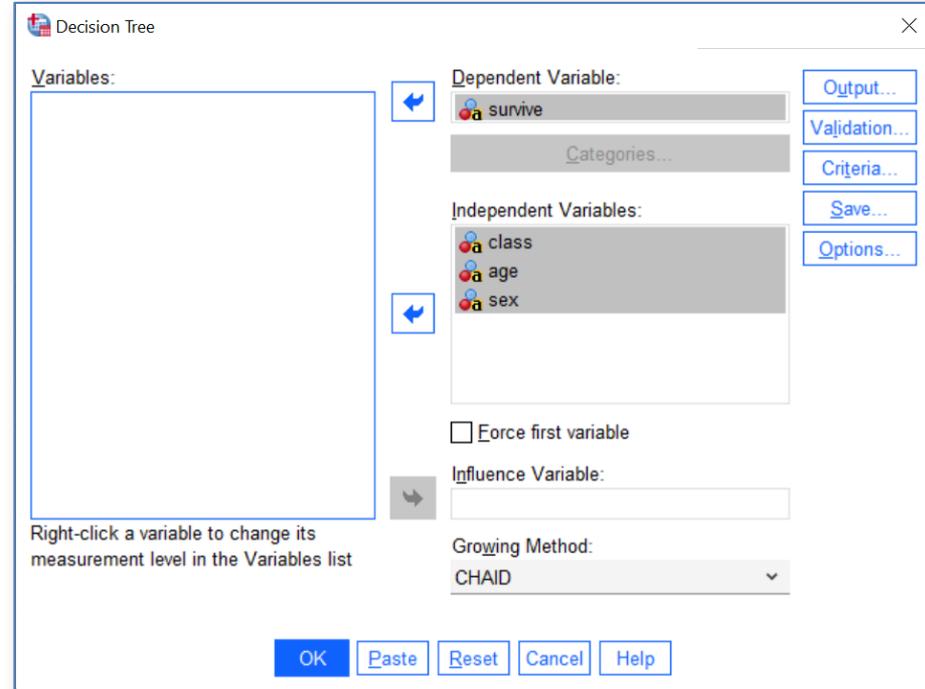




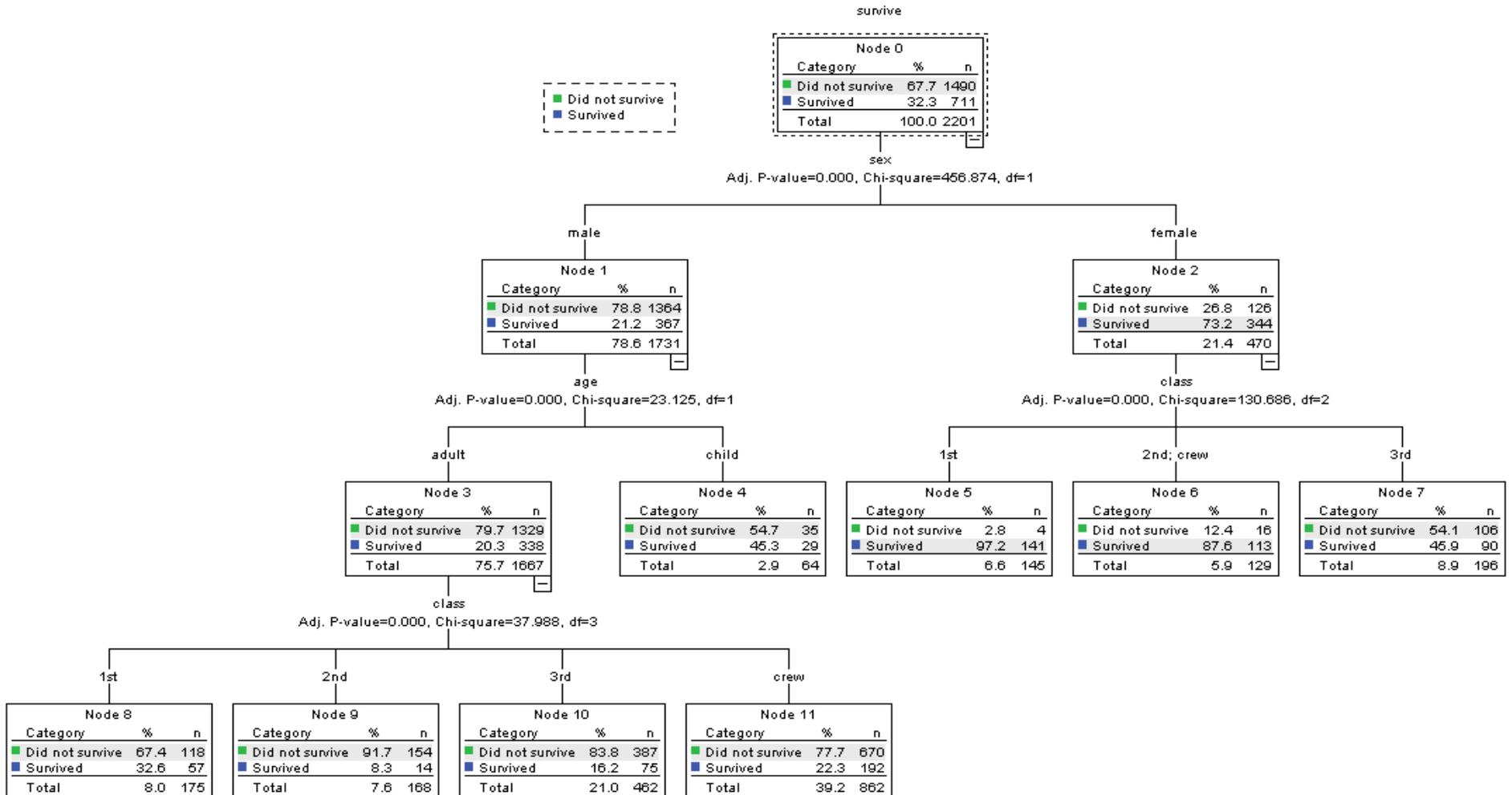
# Example Module: IBM SPSS Decision Trees

# SPSS Decision Trees

- Decision trees are used to:
  - Build predictive models
  - Create profiles of customers/employees/clients
  - Find key behavioural segments
- Very popular because they:
  - Represent models visually
  - Are easy to understand



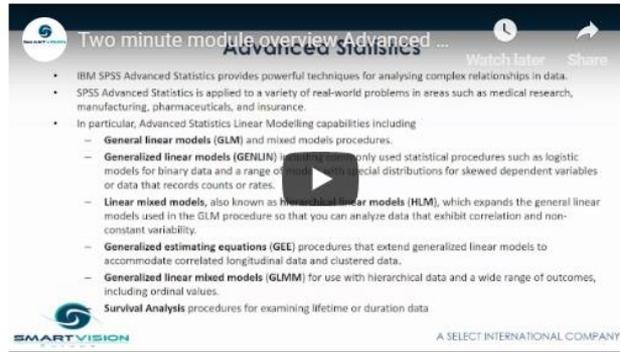
# IBM SPSS Decision Trees



# Smart Vision's Short Video Guides to SPSS Modules

## IBM SPSS Advanced Statistics

### Two minute module overview



## What's in the Advanced Statistics module?

IBM SPSS Advanced Statistics provides powerful techniques for real-world problems in a variety of disciplines, including medical research, manufacturing, pharmaceuticals, and market research. It is an extremely 'analytically rich' SPSS Statistics module. Advanced Statistics includes a very wide range of multivariate procedures for investigating complex relationships in data. A number of the procedures in IBM SPSS Advanced Statistics are relatively technical in a statistical sense. In particular, Advanced Statistics encompasses General and Generalized Linear Modelling capabilities.

- General Linear Models allow you to model relationships and interactions between many factors. The general linear model incorporates a number of different statistical models: ANOVA, MANOVA, ANCOVA, Repeated Measures etc.
- An example of GLM would be one where we wish to see if males spend more money on average in a particular store than females when they have been given a particular discount coupon. The model can be made more complicated if we include more than one type of coupon or test to see whether region or age makes a difference.
- Generalized Linear Models are an extension of General Linear Models in that they are able to work with a greater range of data distributions. In particular, the model allows for the dependent variable to have a non-normal distribution.

## SPSS modules

[IBM® SPSS® Statistics Base](#)

[IBM® SPSS® Advanced Statistics](#)

[IBM® SPSS® Bootstrapping](#)

[IBM® SPSS® Categories](#)

[IBM® SPSS® Complex Samples](#)

[IBM® SPSS® Conjoint](#)

[IBM® SPSS® Custom Tables](#)

[IBM® SPSS® Data Preparation](#)

[IBM® SPSS® Decision Trees](#)

[IBM® SPSS® Direct Marketing](#)

[IBM® SPSS® Exact Tests](#)

[IBM® SPSS® Forecasting](#)

[IBM® SPSS® Missing Values](#)

[IBM® SPSS® Neural Networks](#)

[IBM® SPSS® Regression](#)

[IBM® SPSS® Bundles](#)

# Three Key Capabilities

The most effective users of technology like SPSS Statistics are able to demonstrate skills in three key areas:

1. Reading and formatting data
2. Transforming and manipulating data
3. Analysis and deployment of results



Let's take a closer look

# Additional Resources

- [SPSS FAQs](#) everything from finding out what you have installed to how to merge files or change the language
- [Video Guides](#) a wide range of SPSS “how to” topics with mini demos
- [SPSS Software](#) information on products, modules and pricing
- [Eat your greens](#) blog series on statistical testing and procedures



ChatSPSS is a free generative AI chatbot developed by Smart Vision Europe Ltd based on official IBM documentation of SPSS procedures

The screenshot shows the ChatSPSS website interface. At the top, there is a navigation bar with links for 'NEW CONVERSATION', 'EXPORT PDF', 'LOG OUT', 'PRIVACY', and 'SMARTVISION'. The main content area is divided into two columns. The left column contains a list of 'CONVERSATIONS' and 'DOCUMENT REFERENCES'. The right column displays the chat history and the current response to a user query.

**CONVERSATIONS**

- How can I recode a string variable called region to create a numeric equivalent?
- How can I recode a string variable called region to create a numeric equivalent?

**DOCUMENT REFERENCES**

- IBM\_SPSS\_Statistics\_Algorithms.pdf
- IBM\_SPSS\_Statistics\_Command\_Syntax\_Re...

**Chat History:**

User: How can I recode a string variable called region to create a numeric equivalent?

ChatSPSS: To recode a string variable called `region` to create a numeric equivalent, you can use the `RECODE` command with the `INTO` keyword.

### Steps to Recode the Variable

- Use the `RECODE` command to specify the original values and the new values.
- Use the `INTO` keyword to specify the target numeric variable.
- Ensure that the target variable is declared before using it.

### Example Syntax

```
RECODE region ('North'=1) ('South'=2) ('East'=3) ('West'=4) INTO region_num.
```

- In this example, the string variable `region` is recoded into the numeric variable `region_num`.
- The original values 'North', 'South', 'East', and 'West' are replaced with the numeric values 1, 2, 3, and 4, respectively.

### Alternative Method

Alternatively, you can use the `AUTORECODE` command:

```
AUTORECODE VAI region /INTO region_num /PRINT.
```

# Smart Vision Europe: Services and Expertise

We have decades of experience providing guidance, training and consultancy in the delivery of effective data science initiatives.

# Working with Smart Vision Europe

- **Sourcing Software**
  - You can buy your analytical software from us often with discounts
  - Assist with selection, pilot, implementation & support of analytical tools
  - <http://www.sv-europe.com/buy-spss-online/>
- **Training**
  - Formal classroom/virtual training
  - Online self-paced training resources
- **Advice and Support**
  - ‘No strings attached’ technical and business advice relating to analytics
  - Tracked technical support services around the IBM SPSS product line



Contact us:

+44 (0)207 786 3568

[info@sv-europe.com](mailto:info@sv-europe.com)

Twitter: @sveurope

[Follow us on Linked In](#)



Thank you