DegreeLive

B.COM HONOURS	SEMESTER 3	CALICUT UNIVERSITY	COURSE CODE: COM3MN210

DATA ANALYTICS WITH STATISTICAL SOFTWARE

EXAM REVISION NOTES

JOIN WHATSAPP CHANNEL →

MODULE I

AN INTRODUCTION TO SPSS

Meaning, Evolution, and Nature of SPSS

Applications and Practical Uses of SPSS in Business: In contemporary corporate environments, IBM SPSS serves as a vital decision-support engine across diverse business functions:

- **Market Research Consumer Insights:** ; Identifies target consumer clusters, optimizes product pricing thresholds, and quantifies brand equity drivers.
- **Human Resource Management (HRM):** ; Reduces costly employee turnover, identifies workplace burnout determinants, and benchmarks organizational climate.
- **Banking, Credit Financial Services:** ; Minimizes non-performing assets (NPAs), establishes automated loan approval scorecards, and prevents revenue loss.
- **Retail Supply Chain Management:** ; Enhances inventory turnover ratios, eliminates stockout penalties, and tailors localized promotional strategies.

Step-by-Step Installation Setup Guide for SPSS

Creating, Structuring, and Editing an SPSS Data File (.sav)

Advanced Theoretical Extensions Analytical Frameworks: In modern corporate governance and professional commerce curricula, a mastery of Data Analytics with Statistical Software requires rigorous synthesis of foundational principles with multi-stakeholder operational realities.

Step-by-Step Implementation Roadmap: Executing systemic improvements within Data Analytics with Statistical Software requires an organized four-phase operational roadmap:

- **Phase I: Baseline Assessment amp; Diagnostic Audit:** Phase I: Baseline Assessment amp; Diagnostic Audit: Benchmarking current operating proficiencies against domestic and international industry standards to identify gaps in competence, regulatory readiness, and infrastructure.
- **Phase II: Strategy Formulation amp; Policy Design:** Phase II: Strategy Formulation amp; Policy Design: Formulating standard operating procedures (SOPs), resource allocation schedules, and clear Key Performance Indicators (KPIs) with designated accountability.
- **Phase III: Operational Rollout amp; Change Enablement:** Phase III: Operational Rollout amp; Change Enablement: Executing structured pilot initiatives, cross-functional training seminars, and system migrations with active management coaching.
- **Phase IV: Continuous Governance amp; Iterative Refinement:** Phase IV: Continuous Governance amp; Iterative Refinement: Institutionalizing regular review cycles, variance exception reporting, and multi-stakeholder feedback loops to refine policies iteratively.

Strategic Synthesis Quantitative Performance Analytics: Achieving excellence in Data Analytics with Statistical Software requires an appreciation of quantitative analytical methods and robust performance measurement systems.

Technological Infrastructure Modern Enterprise Stacks: Modern organizations execute functions related to Data Analytics with Statistical Software through sophisticated multi-layered software architectures.

- **Unified Data Ingestion Layer:** Captures transactional data, customer interactions, and audit trails instantaneously from distributed touchpoints, ensuring data integrity at the source.
- **Automated Business Logic amp; Decision Engines:** Automated Business Logic amp; Decision Engines: Applies complex statutory rules, accounting standards, and risk formulas deterministically, reducing manual processing errors and turnaround t × .
- **Executive Visual Dashboards amp; Alert Telemetry:** Executive Visual Dashboards amp; Alert Telemetry: Surfaces actionable exceptions directly to department heads, highlighting anomalous transactions or operational bottlenecks before regulatory thresholds are breached.

Institutional Governance, Statutory Precedents Standard Operating Protocols: To institutionalize sustainable excellence in Data Analytics with Statistical Software, enterprises rely on codified Standard Operating Procedures (SOPs) and comprehensive internal control frameworks.

Capital Allocation Sensitivity Modeling Framework: Deploying major initiatives under Data Analytics with Statistical Software requires continuous modeling against volatile macroeconomic factors:

- **Discounted Cash Flow (DCF) Hurdle Rates:** Capital investment projects are benchmarked against the weighted average cost of capital (WACC) with an added risk premium for macroeconomic uncertainty.
- **Dynamic Break-Even Sensitivity Analysis:** Evaluating operating leverage under conservative, baseline, and optimistic scenarios to ensure solvency during market contractions.
- **Asset-Liability Duration Matching:** Aligning long-term project capital outlays with appropriate non-current funding lines to avert structural liquidity shortages.

Longitudinal Industry Paradigm Shifts: The historical trajectory of Data Analytics with Statistical Software reflects three distinct phases of evolutionary progress:

- **First Generation (Manual amp; Paper-Based Records):** First Generation (Manual amp; Paper-Based Records): Operations were characterized by physical ledgers, localized decision-making, and high latency in transaction clearing.
- **Second Generation (Relational Databases amp; Siloed Automation):** Second Generation (Relational Databases amp; Siloed Automation): Transitioned to desktop computing, functional ERP systems, and batch processing, though functional silos remained pervasive.
- **Third Generation (Cloud-Native, Real-Time Ecosystems):** Current state marked by microservices, algorithmic decision-making, open API connectivity, and continuous statutory telemetry.

International Commercial Protocols Matrix: By mastering cross-border governance mechanisms, professionals managing Data Analytics with Statistical Software enable resilient international commercial expansion and protect enterprise solvency.

- **Settlement Rail Mechanism:** ; Hedging currency exposure through forward contracts and currency options.
- **Contractual Dispute Resolution:** ; Incorporation of standard institutional arbitration clauses.
- **Sanctions amp; AML Screening:** ; Real-time transaction filtering integrated with core banking/ERP interfaces.

MODULE II

INFERENTIAL STATISTICS

Hypothesis Formulation Testing Architecture in SPSS

The Decision Framework: Type I Type II Errors and Significance Levels: Because statistical inference relies on probability rather than absolute certainty, decision-makers are exposed to two fundamental categories of decision error:

- **Reject H sub0;(Claim an effect exists):** Type I Error (α - Alpha) "False Positive" Rejecting a true null hypothesis.
- **Fail to Reject H sub0;(Claim no effect detected):** ; Type II Error (β - Beta) "False Negative" Failing to reject a false null hypothesis.

Mathematical Foundation Calculation of Expected Frequencies: The test compares the Observed Frequencies (O)—the actual case counts tallying in each cell of a cross-tabulation grid—against the Expected Frequencies (E)—the theoretical counts that would occur if the two variables were completely independent.

Interpreting SPSS Chi-Square Output Tables: Examine the table labeled Chi-Square Tests:

- **Pearson Chi-Square:** Locate the top row: .
- **Value:** Inspect the column (chi; sup2; statistic), the df column (degrees of freedom), and the Asymptotic Significance (2-sided) column (p-value).
- **Rule:** If Asymp.
- **Inspect the footnote below the table:** 0%) have expected count less than 5.

Non-Parametric Alternatives: Spearman's Rank Kendall's Tau-b: When research data fails the stringent parametric assumptions of normality, or when data is measured on an ordinal scale (such as Likert satisfaction ratings, rankings, or performance tiers), non-parametric correlation techniques are deployed:

- **Spearman's Rank-Order Correlation (rho; - Rho / r subs;):** Spearman's Rank-Order Correlation (rho; - Rho / r subs;): Computes Pearson's r on the converted rank orders of the data rather than raw scores.
- **Kendall's Tau-b (tau; subb;):** Kendall's Tau-b (tau; subb;): A robust non-parametric measure based on concordant and discordant pairs of observations.

Contemporary Managerial Dilemmas Strategic Solutions: Organizations navigating Data Analytics with Statistical Software must systematically manage complex organizational trade-offs:

- **Operational Efficiency versus Risk Mitigation:** Aggressive cost-cutting can increase vulnerabilities in quality control, statutory compliance, and workforce retention.
- **Centralized Standardization versus Decentralized Agility:** While corporate governance mandates uniform operating procedures, frontline managers need flexibility to respond rapidly to changing customer demands and local market shifts.
- **Digital Automation versus Human Capital Continuity:** Integrating automated analytics platforms requires strategic workforce upskilling to maintain institutional knowledge and morale.

Comprehensive Enterprise Risk Management (ERM) Matrix: Management teams deploy structured risk matrices to categorize, evaluate, and mitigate vulnerabilities across four core exposure categories:

- **Strategic Risk:** ; Continuous scenario modeling, product portfolio diversification, and agile capital budgeting buffers.
- **Operational Risk:** ; Redundant vendor agreements, predictive maintenance algorithms, and rigorous workforce certification.
- **Financial amp; Credit Risk:** ; Dynamic hedging instruments, conservative credit underwriting standards, and dedicated emergency liquidity reserves.
- **Compliance amp; Legal Risk:** ; Dedicated statutory compliance teams, external legal audits, and enterprise-wide whistle-blower mechanisms.

Multi-Tier Governance and Audit Checkpoints: To ensure systemic consistency across decentralized units, enterprises deploy a multi-tier audit schedule designed to detect and remediate process deviations before customer impact occurs:

- **Tier 1: Supervisory Spot-Check:** ; Immediate corrective feedback and on-the-spot procedural adjustments.
- **Tier 2: Functional Process Review:** ; Standard operating procedure updates and targeted employee retraining programs.
- **Tier 3: Executive Governance Audit:** ; Formal board-level audit reports and capital reallocation directives.

Future Frontiers: Five-Year Industry Outlook: Over the coming decade, the domain of Data Analytics with Statistical Software will be profoundly reshaped by three transformative industry vectors:

- **Autonomous Algorithmic Operations:** Routine transactional tasks will be largely automated via intelligent process automation (IPA), freeing commercial professionals to focus on strategic capital allocation, negotiations, and risk stewardship.
- **Stakeholder-Centric ESG Governance:** Capital allocation decisions by institutional funds will increasingly reward organizations demonstrating transparent, verifiable sustainability records alongside healthy financial margins.
- **Cross-Border Regulatory Harmonization:** As digital supply chains and fintech rails interconnect global markets, professionals fluent in both domestic Indian compliance and international standards will enjoy superior strategic relevance.

Strategic Value Modeling, Cost Optimization Capital Stewardship: The long-term value realization in Data Analytics with Statistical Software relies on disciplined capital management, sensitivity analysis, and rigorous cost accounting models.

Longitudinal Industry Paradigm Shifts: The historical trajectory of Data Analytics with Statistical Software reflects three distinct phases of evolutionary progress:

- **First Generation (Manual amp; Paper-Based Records):** First Generation (Manual amp; Paper-Based Records): Operations were characterized by physical ledgers, localized decision-making, and high latency in transaction clearing.
- **Second Generation (Relational Databases amp; Siloed Automation):** Second Generation (Relational Databases amp; Siloed Automation): Transitioned to desktop computing, functional ERP systems, and batch processing, though functional silos remained pervasive.
- **Third Generation (Cloud-Native, Real-Time Ecosystems):** Current state marked by microservices, algorithmic decision-making, open API connectivity, and continuous statutory telemetry.

International Commercial Protocols Matrix: By mastering cross-border governance mechanisms, professionals managing Data Analytics with Statistical Software enable resilient international commercial expansion and protect enterprise solvency.

- **Settlement Rail Mechanism:** ; Hedging currency exposure through forward contracts and currency options.
- **Contractual Dispute Resolution:** ; Incorporation of standard institutional arbitration clauses.
- **Sanctions amp; AML Screening:** ; Real-time transaction filtering integrated with core banking/ERP interfaces.

MODULE III

STATISTICAL ANALYSIS

The Family of Student's t-Tests: Foundations Assumptions

Purpose and Business Applications: The One-Sample t-Test evaluates whether the sample mean (macr;x) of a continuous metric differs significantly from a predetermined, specified benchmark, historical standard, or hypothesized population mean (μ sub0;).

Purpose and Research Design: The Independent-Samples t-Test evaluates whether the population means of two completely separate, independent groups differ significantly on a continuous dependent variable.

- Comparing average monthly customer spend between Prime Members vs.

- Comparing sales productivity between employees trained under Method A vs.
- Evaluating whether website session duration differs between Mobile App users and Desktop Web users.

Purpose and Experimental Design: The Paired-Samples t-Test evaluates mean differences across two related, matched, or repeated observations drawn from the same group of participants across two measurement occasions or experimental conditions.

- **Pre-Test vs. Post-Test Studies:** Assessing employee sales performance before and after a 30-day digital marketing training bootcamp.
- **Matched-Pair Case Control:** Comparing expenditure between two stores matched precisely on physical square footage and neighborhood demographics.
- **Cross-Over Product Trials:** Evaluating customer rating of Product A vs.

Partitioning of Variance The F-Ratio

Factorial Designs Interaction Effects: Two-Way ANOVA (General Linear Model rarr; Univariate in SPSS) evaluates the simultaneous effects of two categorical independent factors (Factor A and Factor B) on a single continuous dependent variable.

- **Main Effect of Factor A:** Does Factor A significantly influence the dependent variable, averaged across all levels of Factor B?
- **Main Effect of Factor B:** Does Factor B significantly influence the dependent variable, averaged across all levels of Factor A?
- **Interaction Effect (A × B):** Does the effect of Factor A depend upon the specific level of Factor B?
- **Factor A:** Locate the rows labeled: , Factor B, and Factor A * Factor B (the interaction).
- **F:** Examine the and Sig.

Contemporary Risk Governance, Statutory Compliance Case Analysis: Corporate decision-makers must ensure that all initiatives within Data Analytics with Statistical Software withstand rigorous scrutiny across legal, financial, and ethical dimensions.

Operational Performance Equations Diagnostics: Analytical rigor in Data Analytics with Statistical Software is reinforced by continuous metric monitoring across operational throughput, variance thresholds, and resource efficiency.

- **Efficiency Variance Ratio:** Measures actual resource consumption against standard engineering benchmarks: Efficiency Variance = (Standard Quantity minus; Actual Quantity) × Standard Rate.
- **Operational Throughput Velocity:** Measures the speed with which inputs are converted into completed outputs across the workflow: Throughput Velocity = Total Processed Units / Total Cycle Time Hours.
- **Total Quality Conformance Index:** The proportion of delivered units meeting strict compliance criteria on initial inspection: Quality Conformance = (Total Conforming Units / Total Units Produced) × 100%.

Global Standards Matrix: Comparative International Analysis: The table below highlights how key functional dimensions of Data Analytics with Statistical Software align across Indian domestic statutory mandates and leading international frameworks:

- **Financial amp; Operational Disclosure:** ; Dual-reporting systems with automated chart-of-accounts translation and real-time reconciliation.
- **Data Security amp; Privacy Compliance:** ; End-to-end data encryption, strict role-based access control (RBAC), and immutable audit logs.
- **Sustainability amp; ESG Reporting:** ; Integrated annual ESG metrics audited by independent certified environmental verifiers.
- **Supply Chain amp; Partner Governance:** ; Automated vendor onboarding, dynamic credit checks, and regular ethical third-party supplier audits.

Statutory Case Precedents and Commercial Jurisprudence: Modern commercial governance has been extensively shaped by landmark legal precedents and statutory rulings. Commercial leaders in Data Analytics with Statistical Software draw crucial guidance from established case law principles:

- **Strict Fiduciary Accountability:** ; Strict Board authorization schedules and explicit signing limit matrices.
- **Doctrine of Constructive Notice:** ; Mandatory legal vetting of non-standard commercial covenants prior to execution.
- **Substance Over Legal Form:** ; Independent compliance sign-off on major commercial agreements.

Longitudinal Sectoral Trends, Technological Evolution Future Architecture: Enterprise success across the spectrum of Data Analytics with Statistical Software is heavily influenced by macroeconomic megatrends, demographic shifts, and evolving consumer expectations.

Multilateral Trade Regulatory Alignment: International commerce in Data Analytics with Statistical Software functions within key global governance architectures:

- **Incoterms 2020 Covenants:** Allocating precise risk boundaries, freight obligations, and insurance responsibilities between contracting parties across maritime and multimodal shipments.
- **Foreign Exchange Management Act (FEMA) Compliance:** Regulating capital and current account transactions, export invoice regularizations, and inward remittance clearances.
- **Base Erosion amp; Profit Shifting (BEPS) Framework:** Base Erosion amp; Profit Shifting (BEPS) Framework: Ensuring arm's-length transfer pricing documentation and economic substance in cross-border intercompany transactions.

MODULE IV

PREDICTIVE MODELING

Meaning, Definition, and Importance of Predictive Modeling

Mapping Business Problems to Analytical Techniques: Selecting the appropriate analytical technique depends upon data scale, distribution, and variable measurement levels:

- **Sales Financial Forecasting:** Continuous Scale (Metric); Multiple Linear Regression, Polynomial Regression, Ridge/Lasso; Analyze rarr; Regression rarr; Linear...
- **Credit Scoring Churn Prediction:** Binary Categorical (0 / 1); Binary Logistic Regression, Linear Discriminant Analysis (LDA); Analyze rarr; Regression rarr; Binary Logistic...
- **Customer Preference Multi-Class:** Multinomial Nominal (3+ classes); Multinomial Logistic Regression, Decision Trees (CHAID, CART); Analyze rarr; Regression rarr; Multinomial Logistic...
- **Market Segmentation:** None (Unsupervised); K-Means Cluster Analysis, Two-Step Clustering, Hierarchical Clustering; Analyze rarr; Classify rarr; K-Means Cluster...

Key Data Preparation Protocols in SPSS

- **Dummy Variable Encoding:** Statistical regression models require numerical inputs.
- **Train-Test Holdout Partitioning:** To prevent statistical overfitting (where a model memorizes sample noise rather than generalizable population relationships), the dataset is partitioned into an estimation/training sample (typically 70-80%) and an independent holdout test sample (20-30%) via Transform rarr; Random Number Generators.

DegreeLive

Outlier Detection and Treatment: Outliers are extreme observations that deviate so markedly from the general distribution that they distort parameter estimates, inflate standard errors, and bias regression coefficients:

- **Univariate Outlier Detection:** Calculated using standardized Z-scores.
- **Multivariate Outlier Detection:** Evaluated using Mahalanobis Distance (D sup2;), which measures the multi-dimensional distance of a case from the centroid of all predictor variables.
- **Treatment Options:** , typing 250 instead of 25 for age).
- **Winsorization (Capping):** Replace extreme outliers with the value of the 95th or 99th percentile threshold.
- **Mathematical Transformation:** Applying logarithmic or square-root transformations to compress long right-hand tails in positively skewed financial metrics.

Essential Pre-Requisites Diagnostic Tests: Before extracting factors in SPSS, the analyst must verify that the correlation matrix is suitable for factor analysis:

- **Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy:** Evaluates the proportion of variance among variables that might be common variance.

Advanced Theoretical Extensions Analytical Frameworks: In modern corporate governance and professional commerce curricula, a mastery of Data Analytics with Statistical Software requires rigorous synthesis of foundational principles with multi-stakeholder operational realities.

Contemporary Risk Governance, Statutory Compliance Case Analysis: Corporate decision-makers must ensure that all initiatives within Data Analytics with Statistical Software withstand rigorous scrutiny across legal, financial, and ethical dimensions.

Operational Performance Equations Diagnostics: Analytical rigor in Data Analytics with Statistical Software is reinforced by continuous metric monitoring across operational throughput, variance thresholds, and resource efficiency.

- **Efficiency Variance Ratio:** Measures actual resource consumption against standard engineering benchmarks: Efficiency Variance = (Standard Quantity minus; Actual Quantity) × Standard Rate.
- **Operational Throughput Velocity:** Measures the speed with which inputs are converted into completed outputs across the workflow: Throughput Velocity = Total Processed Units / Total Cycle Time Hours.
- **Total Quality Conformance Index:** The proportion of delivered units meeting strict compliance criteria on initial inspection: Quality Conformance = (Total Conforming Units / Total Units Produced) × 100%.

Global Standards Matrix: Comparative International Analysis: The table below highlights how key functional dimensions of Data Analytics with Statistical Software align across Indian domestic statutory mandates and leading international frameworks:

- **Financial amp; Operational Disclosure:** ; Dual-reporting systems with automated chart-of-accounts translation and real-time reconciliation.
- **Data Security amp; Privacy Compliance:** ; End-to-end data encryption, strict role-based access control (RBAC), and immutable audit logs.
- **Sustainability amp; ESG Reporting:** ; Integrated annual ESG metrics audited by independent certified environmental verifiers.
- **Supply Chain amp; Partner Governance:** ; Automated vendor onboarding, dynamic credit checks, and regular ethical third-party supplier audits.

Statutory Case Precedents and Commercial Jurisprudence: Modern commercial governance has been extensively shaped by landmark legal precedents and statutory rulings. Commercial leaders in Data Analytics with Statistical Software draw crucial guidance from established case law principles:

- **Strict Fiduciary Accountability:** ; Strict Board authorization schedules and explicit signing limit matrices.
- **Doctrine of Constructive Notice:** ; Mandatory legal vetting of non-standard commercial covenants prior to execution.
- **Substance Over Legal Form:** ; Independent compliance sign-off on major commercial agreements.

Longitudinal Sectoral Trends, Technological Evolution Future Architecture: Enterprise success across the spectrum of Data Analytics with Statistical Software is heavily influenced by macroeconomic megatrends, demographic shifts, and evolving consumer expectations.

Multilateral Trade Regulatory Alignment: International commerce in Data Analytics with Statistical Software functions within key global governance architectures:

- **Incoterms 2020 Covenants:** Allocating precise risk boundaries, freight obligations, and insurance responsibilities between contracting parties across maritime and multimodal shipments.
- **Foreign Exchange Management Act (FEMA) Compliance:** Regulating capital and current account transactions, export invoice regularizations, and inward remittance clearances.
- **Base Erosion amp; Profit Shifting (BEPS) Framework:** Base Erosion amp; Profit Shifting (BEPS) Framework: Ensuring arm's-length transfer pricing documentation and economic substance in cross-border intercompany transactions.

DegreeLive