Collaborative Business Case - Factor Analysis

Customer Satisfaction Analysis: TechnoServe Solutions

Modality: Collaborative (teams of 2-3 students) **Weight:** 5% of total course grade **Duration:** 1 week **Deliverables:** Jupyter Notebook + Executive Summary + Presentation Video (YouTube)

Problem Context

You are part of a data science consulting team tasked with helping **TechnoServe Solutions**, a technology consulting firm, understand their customer satisfaction survey data using factor analysis.

Business Situation: TechnoServe has collected customer satisfaction data but needs to identify the key underlying factors that drive customer retention and business growth.

Your Mission: Apply Factor Analysis to discover meaningful patterns in satisfaction data and present actionable insights to the executive team.

Learning Objectives

Upon completing this business case, students will be able to:

- **Technical:** Implement basic Factor Analysis using Python with proper interpretation
- **Methodological:** Assess data suitability and determine appropriate number of factors
- **Analytical:** Interpret factor loadings and give factors meaningful business names
- **Professional:** Communicate statistical findings to business audiences clearly
- **Collaborative:** Work effectively in teams to solve data analysis problems

---

Dataset and Business Context

Customer Satisfaction Dataset

Source: Customer Satisfaction Survey (2024)

- **Observations:** 3,400 survey responses
- **Clients:** 850 enterprise customers
- **Variables:** 23 satisfaction dimensions (1-7 scale) + 5 outcome variables
- **Context:** Technology consulting firm with multiple service areas

Satisfaction Variables (23 dimensions)

Technical Excellence & Innovation:

- `technical_expertise`, `problem_solving`, `innovation_solutions`, `technical_documentation`, `system_integration`

Relationship Management & Communication:

- `account_manager_responsive`, `executive_access`, `trust_reliability`, `long_term_partnership`, `communication_clarity`

Project Delivery & Quality:

- `project_management`, `timeline_adherence`, `budget_control`, `quality_deliverables`, `change_management`

Value & Financial Transparency:

- `cost_transparency`, `value_for_money`, `roi_demonstration`, `competitive_pricing`, `billing_accuracy`

Support & Service Excellence:

- `support_responsiveness`, `training_quality`, `documentation_help`

Outcome Variables (for validation)

- **`overall_satisfaction`** (1-7): Overall satisfaction with TechnoServe Solutions
- **`nps_score`** (0-10): Net Promoter Score (likelihood to recommend)
- **`renewal_likelihood`** (1-5): Contract renewal probability
- **`revenue_growth_pct`** (continuous): Year-over-year revenue growth percentage
- **`referrals_generated`** (integer): Number of referrals generated per quarter

---

Simplified Project Structure

Part 1: Data Exploration and Suitability (25 points)

1.1 Basic Data Exploration (10 points)

Tasks:

- Load and examine the dataset structure
- Calculate basic descriptive statistics
- Create correlation matrix visualization
- Identify patterns in the data

Deliverable: Clear summary of data characteristics and patterns

1.2 Factor Analysis Suitability (15 points)

Required Tests:

- **KMO Test:** Overall sampling adequacy (should be > 0.6)
- **Correlation Assessment:** Check if variables are sufficiently correlated
- **Basic Assumptions:** Evaluate if factor analysis is appropriate

Decision Criteria:

- KMO > 0.6 (acceptable for factor analysis)
- Sufficient correlations between variables (>30% with $|r| \geq 0.3$)

Questions to Answer:

- Is the data suitable for factor analysis?
- What do the initial patterns suggest about underlying factors?

Part 2: Factor Extraction and Determination (30 points)

2.1 Determining Number of Factors (15 points)

Methods to Use:

- **Kaiser Criterion:** Eigenvalues > 1.0
- **Scree Plot:** Visual identification of "elbow"
- **Variance Explained:** Cumulative variance analysis

Task: Determine the optimal number of factors and justify your choice

2.2 Factor Extraction and Rotation (15 points)

Implementation:

- Extract factors using Principal Component method
- Apply Varimax rotation for interpretability
- Examine factor loadings matrix

Analysis: Compare factor solutions and select the most interpretable one

Questions to Answer:

- How many factors best represent the data?
- What does each factor represent in business terms?

Part 3: Interpretation and Business Application (30 points)

3.1 Factor Interpretation (15 points)

Factor Labeling:

- Identify variables with high loadings (> |0.4|) on each factor
- Create meaningful business labels for each factor
- Explain what each factor represents for TechnoServe Solutions

Validation:

- Check if factors make business sense
- Evaluate factor solution quality

3.2 Business Insights and Recommendations (15 points)

Factor Scores:

- Calculate factor scores for customers
- Use factor scores to predict outcome variables
- Identify which factors are most important for business outcomes

Strategic Recommendations:

- Prioritize factors based on business impact
- Suggest specific improvement strategies
- Propose action plan for TechnoServe Solutions

Questions to Answer:

- Which factors drive customer satisfaction most?
- What specific actions should TechnoServe take?

Part 4: Communication and Presentation (15 points)

4.1 Visualization (8 points)

Required Plots:

- Factor loadings visualization
- Scree plot for factor selection
- Factor scores distribution
- Business impact summary chart

Quality Standards:

- Clear, professional visualizations
- Appropriate for business audience
- Include explanatory text

4.2 Executive Summary (7 points)

Content Requirements:

- Key findings summary
- Factor interpretations in business language
- Top 3 recommendations for action
- Expected business impact

Format: Professional presentation suitable for executives

Deliverables and Assessment

Deliverable 1: Analysis Notebook (`customer_satisfaction_analysis_team[X].ipynb`)

Required Content:

- Complete factor analysis implementation in Python
- Clear documentation of each step
- Answers to all embedded questions
- Professional code with comments

Deliverable 2: Executive Summary (`executive_summary_team[X].pdf`)

Specifications:

- **Length:** Maximum 2 pages
- **Audience:** TechnoServe Solutions management team
- **Format:** Executive summary with key findings and top 3 recommendations
- **Focus:** Business insights and actionable recommendations

Deliverable 3: Video Presentation (YouTube)

Specifications:

- **Duration:** 8-10 minutes
- **Format:** Professional presentation of findings
- **Participation:** Each team member presents part of the analysis
- **Quality:** Clear audio/video with slides

Suggested Structure:

- **Introduction** (1-2 min): Problem and approach
- **Key Findings** (4-5 min): Factor analysis results and interpretation
- **Recommendations** (2-3 min): Business recommendations and next steps

Team Information Template

In final cell of notebook:

```
## Team Information

**Team:** [Team name]

**Members:**
- [Full Name 1] ([ID]) - Data exploration and factor extraction
- [Full Name 2] ([ID]) - Factor interpretation and business insights
- [Full Name 3] ([ID]) - Visualization and recommendations

**Deliverable Links:**
- **Presentation Video:** [YouTube Link]
- **Executive Summary:** [Available on Canvas]
- **Dataset:** `customer_satisfaction_data.csv`

**Completion Date:** [DD/MM/YYYY]
```

Evaluation Rubric (100 points)

Distribution: 60% Technical Rigor + 25% Business Application + 15% Communication

Component	Points	Assessment Criteria
Data Exploration and Suitability	25	EDA (8) + Statistical tests (10) + Correlation analysis (7)
Factor Extraction and Determination	30	Number determination (15) + Extraction and rotation (15)
Business Interpretation	25	Factor labeling (12) + Business insights and recommendations (13)
Communication and Presentation	20	Visualization (12) + Executive summary and video (8)

Detailed Rubric by Competency

1. Technical and Methodological Rigor (55 points)

Level	Outstanding	Competent	Sufficient	Insufficient	Not Submitted
Data Exploration & Suitability (25 pts)	**25-23 pts:** Comprehensive EDA, proper KMO interpretation, excellent correlation analysis	**22-20 pts:** Good exploration, correct tests, appropriate interpretation	**19-17 pts:** Basic exploration, tests present, minor errors	**16-10 pts:** Partial implementation, interpretation errors	**<10 pts:** Tests absent or incorrect

Level	Outstanding	Competent	Sufficient	Insufficient	Not Submitted
Factor Extraction & Determination (30 pts)	**30-27 pts:** Excellent method comparison, well-justified factor selection, proper rotation	**26-24 pts:** Good implementation, appropriate selection, correct rotation	**23-21 pts:** Basic implementation, reasonable selection	**20-12 pts:** Methodological errors, poor justification	**<12 pts:** Methods absent or incorrect

2. Business Application (25 points)

Level	Outstanding	Competent	Sufficient	Insufficient	Not Submitted
Factor Interpretation (12 pts)	**12-11 pts:** Excellent business labels, strong theoretical coherence, sophisticated analysis	**10-9 pts:** Good labels, solid coherence, clear interpretation	**8-7 pts:** Basic labels, acceptable coherence	**6-4 pts:** Weak labels, poor coherence	**<4 pts:** Interpretation absent
Business Insights & Recommendations (13 pts)	**13-12 pts:** Strategic insights, actionable recommendations, clear impact analysis	**11-10 pts:** Good insights, practical recommendations	**9-8 pts:** Basic insights, simple recommendations	**7-5 pts:** Weak insights, vague recommendations	**<5 pts:** Insights absent

3. Communication and Presentation (20 points)

Level	Outstanding	Competent	Sufficient	Insufficient	Not Submitted
Visualization (12 pts)	**12-11 pts:** Professional graphics, clear insights, effective storytelling	**10-9 pts:** Good visualization, clear message	**8-7 pts:** Functional graphics, acceptable quality	**6-4 pts:** Basic visualization, weak message	**<4 pts:** Visualization absent
Executive Summary & Video (8 pts)	**8-7 pts:** Excellent presentation, balanced participation, professional quality	**6 pts:** Good presentation, appropriate participation	**5 pts:** Functional presentation, acceptable quality	**4-2 pts:** Weak presentation, participation issues	**<2 pts:** Video absent

Additional Excellence Criteria

Innovation Bonus (up to +5 points)

- Implementation of advanced methods not covered in class
- Temporal stability analysis in longitudinal data
- Creative integration with machine learning techniques
- Development of interactive tools for stakeholders

Academic Penalties

- **-10 points:** Academic integrity violations (plagiarism, copied code)
- **-5 points:** Video exceeds 15 minutes or less than 8 minutes
- **-3 points:** Unequal participation in video (>1 minute difference between members)
- **-5 points:** Notebook doesn't execute completely due to code errors

---

Realistic Schedule and Work Plan

One-Week Timeline

Day	Activities	Time Investment
Day 1-2	Team formation, data exploration, suitability tests	2-3 hours total
Day 3-4	Factor determination, extraction, interpretation	3-4 hours total
Day 5-6	Business analysis, visualizations, executive summary	2-3 hours total
Day 7	Video recording, final review, submission	1-2 hours total

Total Team Investment: 8-12 hours over 7 days **Individual Commitment:** 3-4 hours total per person

Daily Work Suggestions

Individual Work (2 hours/day max):

- Code development and analysis
- Individual sections of report
- Preparation for team meetings

Team Coordination (30 minutes/day):

- Daily check-ins (online)
- Progress review
- Division of remaining work

Submission Requirements

- **Deadline:** [To be defined by instructor]
- **Method:** Submit via course platform
- **Files:** Jupyter notebook (.ipynb) + Executive summary (.pdf) + YouTube video link

---

Resources and Support

Required Libraries

```
# Essential libraries only
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from factor_analyzer import FactorAnalyzer, calculate_kmo
from sklearn.preprocessing import StandardScaler
```

Template Structure

```
# Step 1: Data exploration
satisfaction_data.describe()

# Step 2: Check suitability
kmo_all, kmo_model = calculate_kmo(satisfaction_data)

# Step 3: Determine factors
fa = FactorAnalyzer(n_factors=5, rotation='varimax')
fa.fit(satisfaction_data)

# Step 4: Interpret results
loadings = fa.loadings_
```

Getting Started

Required Software

- Python 3.10+ with pandas, numpy, matplotlib, seaborn
- factor_analyzer package for factor analysis
- scikit-learn for additional statistical tools
- Jupyter notebook for analysis

Dataset Information

- **File:** `customer_satisfaction_data.csv`
- **Size:** 3,400 survey responses from 850 customers
- **Variables:** 23 satisfaction items + 5 business outcomes
- **Documentation:** See `CUSTOMER_SATISFACTION_DATA_DICTIONARY.md`

Getting Help

- Consult course materials on factor analysis methodology
- Use help documentation for Python packages
- Ask questions during office hours
- Collaborate with team members on interpretation

Tips for Success

Technical Analysis

- Start with basic data exploration before factor analysis
- Pay attention to KMO test results - they indicate data suitability
- Choose number of factors based on multiple criteria (eigenvalues, scree plot, business logic)
- Make sure factor interpretations make business sense

Business Application

- Think about what each factor represents for TechnoServe's business
- Focus on actionable insights rather than just statistical results
- Connect factor analysis results to business outcomes
- Provide specific, practical recommendations

Team Collaboration

- Divide work fairly among team members
- Regular check-ins to ensure consistent approach
- Practice presentation together before recording
- Each member should understand all parts of the analysis

This business case integrates advanced statistical rigor, meaningful business application, and professional communication skills development, preparing students for senior roles in analytics and data consulting.