VALIDATING AN EMPLOYEE SATISFACTION INDEX USING CONFIRMATORY FACTOR ANALYSIS IN PYTHON FOR A U.S. TECH COMPANY

1. Background

A mid-sized technology firm based in California conducted an annual engagement survey as part of its internal HR metrics. The survey measured various aspects of employee satisfaction through multiple Likert-scale questions grouped under leadership, work-life balance, growth, and workplace support.

The leadership team wanted to statistically validate whether these survey items genuinely represented four latent constructs. We were brought in to perform Confirmatory Factor Analysis (CFA) in Python and confirm the structural integrity of the proposed satisfaction index.

2. Objective

- To validate the internal employee satisfaction index through CFA in Python
- To determine whether observed survey items reliably loaded onto the four predefined factors
- To assess the model's goodness-of-fit and refine the scale structure as needed
- To support HR with a reliable satisfaction framework for longitudinal tracking

3. Data Used

Source: Internal employee engagement survey (latest cycle)

Dataset Details:

- 947 employees
- 20 Likert-scale items (1 to 5)
- Pre-assigned factors by HR:
    - Leadership (5 items)
    - Work-Life Balance (5 items)
    - Career Growth (5 items)
    - Workplace Support (5 items)

Format: Structured CSV, no missing values

4. Methodology

4.1 Data Preparation

- Checked item distributions and correlations using pandas and seaborn
- Verified suitability for CFA using:
    - **KMO = 0.90 (overall)**
    - **Bartlett's Test**: $p < 0.001$

4.2 Confirmatory Factor Analysis

- Used semopy and pandas for CFA modeling
- Defined a measurement model:
    - 4 latent variables (Leadership, Balance, Growth, Support)
    - Each latent variable mapped to 5 observed variables
- Fit indices used:
    - **CFI**, **TLI**, **RMSEA**, **SRMR**

4.3 Model Evaluation

- Acceptable thresholds:
    - CFI/TLI ≥ 0.90
    - RMSEA ≤ 0.08
    - SRMR ≤ 0.08

5. Factor Structure Results

Latent Factor	Item Loadings Range	Cronbach's Alpha	Notes
Leadership	0.68 – 0.83	0.87	Strong internal consistency
Work-Life Balance	0.61 – 0.79	0.84	Slight cross-loading adjusted

Career Growth	0.66 – 0.85	0.88	Highly distinct and well-loaded
Workplace Support	0.71 – 0.81	0.85	Clean structure, no drops required

Fit indices:

- CFI = 0.93
- TLI = 0.91
- RMSEA = 0.06
- SRMR = 0.05

Model was deemed an excellent fit.

6. Interpretation and Use

- Each factor demonstrated **strong construct validity and internal reliability**
- Suggested minor rewording for 2 items to improve future model clarity
- Recommended:
    - Reporting factor scores per department
    - Tracking latent scores over time instead of just item averages
    - Using model to assess policy impacts (e.g., new hybrid work guidelines)

7. Reporting Output

- **Python Script (semopy + matplotlib)**:
    - Model definition, fit, and output
    - Visual diagram of factor structure
    - Exported standardized loadings and fit statistics
- **PDF Report (17 pages)**:
    - Model diagram
    - Fit index summary
    - Tables: item-to-factor mapping, reliability scores

    - Actionable HR recommendations
- **Excel Sheet**:
    - Respondent-level factor scores
    - Summary by team, department, and office location
    - Score tracker template for future survey cycles

8. Organizational Impact

- HR transitioned to **factor-based satisfaction reporting** for board presentations
- Survey credibility improved, increasing participation rate by 12% in the next cycle
- Factor scores now used as key metrics in **quarterly people analytics reviews**
- CFA output integrated into a Tableau dashboard to monitor engagement trends