

R Analytics Module 2: carrier Hub (360° Partner Matrix)

1. Technical Significance

The **Carrier Hub** (or 360° Partner Matrix) moves beyond simple weighted averages (e.g., “50% price, 50% speed”). It uses **Structural Equation Modeling (SEM)** and **Factor Analysis** to discover hidden (latent) qualities of a carrier that aren’t directly measurable, such as “Reliability” or “Operational Efficiency,” based on observed variables like OTD, Damage Rate, and billing accuracy.

2. The R Technique: SEM + Factor Analysis

We use a psychometric approach to grading carriers:

A. Confirmatory Factor Analysis (CFA)

- **Purpose:** To group related metrics into latent factors.
- **Algorithm:** `psych` and `lavaan` packages in R.
- **Mechanism:**
 - *Observed Variables:* On-Time Delivery (OTD), Damage Free %, Invoice Accuracy, Tracking %
 - *Latent Factors:*
 - * **Reliability:** (correlated with OTD + Tracking)
 - * **Quality:** (correlated with Damage Free + Invoice Acc)
- **Why it’s superior:** It automatically determines *which* metrics matter most. If OTD and Tracking are 99% correlated, the model knows they represent the same underlying trait (“Reliability”) and weights them appropriately, avoiding double-counting.

B. Structural Equation Modeling (SEM)

- **Purpose:** To produce a final, scientific “Carrier Score” (0-100).
- **Algorithm:** `lavaan` package.
- **Mechanism:** It fits a model where the Latent Factors predict the “Overall Performance” construct. The regression weights from this model become the weights for the final scorecard.

3. Workflow & Architecture

1. Data Ingestion:

- The system collects performance metrics for all enabled carriers.
- Request sent to `backend/r_analytics_service.py`.

2. R Execution (`scoring.R`):

- Python passes the carrier data matrix (Rows = Carriers, Cols = Metrics) to R.

- R runs `score_carriers()`.
3. **Statistical Processing:**
- R performs a Principal Component Analysis (PCA) or Factor Analysis.
 - It extracts the “Factor Loadings” (how strongly each metric defines success).
 - It computes a standardized “Factor Score” for each carrier.
4. **Output:**
- R returns a map of `carrier_id -> score` (0-100) and `grade` (A+, B, C).

4. Sample Data & Results

Input Data (Sample)

```
[
  { "id": "CARR_001", "otd": 98, "damage_free": 99, "cost_index": 85 },
  { "id": "CARR_002", "otd": 72, "damage_free": 95, "cost_index": 92 },
  ...
]
```

R Processing Output (Simulated)

```
{
  "scores": {
    "CARR_001": 94.5,
    "CARR_002": 76.2
  },
  "grades": {
    "CARR_001": "A+",
    "CARR_002": "C"
  },
  "factor_loadings": {
    "otd": 0.85,      // OTD is highly predictive of quality
    "cost_index": 0.30 // Cost is less predictive of quality
  }
}
```

User Facing Result

- **TCI Express:** Score **94 (Platinum)** - “Excellent reliability, despite higher cost.”
- **VRL Logistics:** Score **88 (Gold)** - “Good performer, opportunities in billing accuracy.”