QuantSpace Analytics

JSE Concentration Risk Engine

Complete User Guide & Tutorial

For Asset Managers, Wealth Managers, Financial Advisers,

Portfolio Managers and Institutional Investors

Version 3.0 | March 2026

Supports simulated data and licensed JSE constituent uploads

1. What Is the JSE Concentration Risk Engine?

The QuantSpace JSE Concentration Risk Engine is an institutional-grade analytics platform built specifically for the South African equity market. It measures, visualises, and explains concentration risk in the two primary JSE equity benchmarks — the ALSI (FTSE/JSE All Share Index) and the SWIX (Shareholder Weighted Index).

The engine answers the most important and least visible question in South African portfolio management:

> *Are we truly diversified — or are we structurally dependent on a small number of dominant stocks?*

The JSE is one of the most concentrated equity markets in the world. A handful of large companies — Naspers, Prosus, Anglo American, BHP, Richemont — can account for nearly half of the entire ALSI. A portfolio that appears diversified across 40 names may, in practice, be almost entirely driven by five.

This engine makes that concentration visible, measurable, and actionable — for portfolio managers, risk officers, wealth managers, financial advisers, pension fund trustees, and institutional investors.

Version 3.0 introduces a second data mode alongside the simulation: licensed users can now upload official JSE constituent data directly from their own data subscription, custodian, or broker feed, and run the full analytics stack against their actual benchmark data.

> *The simulation mode uses JSE-realistic data for demonstration and workflow validation. It is not connected to live JSE or FTSE Russell data feeds. The Upload JSE Data mode is designed for production use with licensed constituent data only.*

2. Who Benefits From This System?

2.1 Chief Investment Officers and Portfolio Managers

Portfolio managers use the engine to determine whether their active positions are genuinely different from the benchmark, or whether they are replicating benchmark concentration in a different wrapper.

- Is my alpha genuine, or am I just holding Naspers overweight?
- How does my portfolio concentration compare to the ALSI or SWIX?
- Which names drive my benchmark-relative risk?
- What happens to my portfolio if Naspers falls 20% tomorrow?

Upload your licensed JSE constituent file to run all analytics against actual benchmark weights rather than the simulation. The shock scenarios, portfolio comparison, and Excel export all update automatically.

2.2 Risk Officers and Risk Committees

- The Concentration Risk Score provides a single defensible number for board packs and risk dashboards
- The 24-month historical trend shows whether concentration is rising or falling
- The HHI is a regulator-recognised measure of market structure
- Scenario analysis supports ICAAP-style stress testing and Reg 28 reporting conversations
- The Excel export can be attached to meeting packs without further formatting
- Upload official JSE data to ensure the report reflects actual benchmark weights on the reporting date

2.3 Wealth Managers and Financial Planners

- Explain to clients why their JSE equity exposure carries hidden single-stock risk
- Compare a proposed model portfolio against the benchmark to demonstrate genuine diversification
- Show the rand impact of a Naspers drawdown on the client's equity allocation
- Use simulated mode for client demonstrations; switch to uploaded JSE data for client reporting

2.4 Financial Advisers and IFAs

- Use the Risk Score as a client-friendly metric — Low, Moderate, High, Critical
- Show how a portfolio change reduces or increases concentration risk
- Support retirement annuity and living annuity conversations with diversification evidence
- Demonstrate analytical rigour beyond past-return analysis

2.5 Pension Funds and Trustees

Trustees have a fiduciary duty to understand concentration risk embedded in their equity mandates.

- Upload the official JSE constituent file from your data provider to run analytics against actual index weights
- Assess whether a passive ALSI mandate creates unacceptable single-name exposure
- Support Regulation 28 conversations with quantitative concentration evidence
- Request the Excel report as part of quarterly manager reporting requirements

2.6 Data Providers and Custodians

Custodians and data providers who supply JSE constituent files to clients can use the engine to validate that their data exports are correctly formatted and analytically coherent.

- Upload the constituent CSV and confirm the engine accepts and normalises it correctly
- Use the Holdings page to verify constituent counts, sector groupings, and cumulative weights
- Confirm the as-of date is correctly read from the date column in the uploaded file

3. Data Source Modes

Version 3.0 supports two data source modes, selectable from the sidebar. The mode affects what data the benchmark analytics are based on — all other engine behaviour is identical.

	Simulated Data	Upload JSE Data
Best for	Demos, development, workflow testing	Production reporting with licensed data
Data source	Hardcoded JSE-realistic weights in simulate_jse.py	CSV uploaded by the user from their data subscription
Benchmark	ALSI or SWIX via sidebar selector	Whatever is in the uploaded file
History	24-month simulated drift (deterministic, seed 42)	Single as-of snapshot unless historical CSV is uploaded
As-of date	Fixed: December 2025	Read from date column in file, or current run date
Reproducibility	Fully deterministic — identical every run	Depends on the uploaded file contents
Disclaimer required	Yes — simulated data watermark shown	No — replaced by as-of date from file

The analytics engine, metrics module, shock simulator, portfolio comparison, and Excel export are identical in both modes. Only the data feeding into them changes.

4. Key Concepts Explained Simply

4.1 Index Weight

Every stock in the ALSI has a weight — the percentage of the total index it represents. If Naspers has a weight of 21%, every R100 invested in an ALSI tracker puts R21 into Naspers. A 10% fall in Naspers reduces the tracker by approximately 2.1%, regardless of what the other 39 stocks do.

4.2 Top-N Weight

The engine reports the combined weight of the top 1, 3, 5, and 10 stocks.

> *Rule of thumb: if the top 10 stocks account for more than 60% of the index, concentration risk is material and warrants active monitoring.*

4.3 Herfindahl-Hirschman Index (HHI)

The HHI is a well-established measure of market concentration used by economists, regulators, and central banks worldwide. It is the sum of the squares of each constituent's weight expressed as a decimal.

HHI Range	Interpretation
Below 0.01	Highly diversified. Near-equal weighting.
0.01 to 0.05	Moderate concentration. Acceptable for most mandates.
0.05 to 0.10	High concentration. Warrants monitoring and reporting.
Above 0.10	Very high concentration. Material single-name event risk.

4.4 Effective Number of Constituents

> *If the ALSI has an Effective N of 14, the 40-stock index behaves as if it only holds 14 equally weighted stocks. A 40-stock index with Effective N of 14 is less diversified than a 20-stock index with Effective N of 18.*

4.5 Composite Risk Score

A proprietary QuantSpace metric from 0 to 100, combining three pillars:

Pillar	Weight	Danger Threshold
Top-5 concentration	40%	Scaled against a 40% top-5 danger threshold
HHI concentration	40%	Scaled from equal-weight HHI toward a 0.10 ceiling
Effective-N ratio	20%	Gap between Effective N and total constituent count

Classification	Score	What It Means

Low	0 to 25	Index is well diversified. Single-name risk is low.
Moderate	25 to 50	Concentration present. Monitor dominant names.
High	50 to 75	Significant concentration. Active risk management warranted.
Critical	75 to 100	Extreme concentration. Material single-name event risk.

5. Uploading Licensed JSE Constituent Data

This section covers everything you need to know to upload your own licensed JSE constituent data and run the full analytics stack against it.

5.1 When to Use Upload Mode

- You hold a JSE direct data subscription, FTSE Russell license, broker data feed, or custodian data export
- You are preparing a client report or board pack that must reflect official benchmark weights
- You want to run concentration analytics against a non-standard benchmark or a custom universe
- You are validating that your data provider's export format is analytically correct

Do not use simulated data mode for compliance reporting, client reporting, or production risk management. Switch to Upload JSE Data mode and use your licensed constituent file.

5.2 Required File Format — Current Snapshot

The upload expects a CSV file with four required columns:

Column	Required	Type	Description
ticker	Yes	string	Instrument or constituent code. Uppercased and stripped of whitespace automatically.
name	Yes	string	Constituent display name. Blank names fall back to the ticker value.
sector	Yes	string	Sector classification for aggregation and charts. Blank sectors become Unknown.
weight	Yes	numeric	Constituent weight. May be percentage, fractional, or raw market-cap value. Normalised to 100 automatically.

Example file content:

```
ticker,name,sector,weight
NPN,Naspers Ltd,Technology,10.5
SBK,Standard Bank Group,Financials,5.6
BHP,BHP Group plc,Mining,3.1
```

5.3 Optional Date Column

If your data provider includes a snapshot date in the file, the engine reads it automatically. Any of the following column names are accepted:

- date
- as_of_date
- effective_date

The engine uses the latest valid date value found in the column as the uploaded data as-of date. This date appears in the sidebar, the Dashboard header, and the Excel report cover sheet. If no date column is present, the engine uses the current run date.

5.4 Upload Validation Rules

The engine applies the following validation and cleaning steps in order when a constituent file is uploaded:

1. Files missing any of ticker, name, sector, or weight columns are rejected with an error message.
2. Column names are normalised to lowercase and stripped of whitespace.
3. Ticker values are converted to uppercase and stripped of whitespace.
4. Weight values are coerced to numeric. Non-numeric values are dropped with a warning.
5. Rows with blank tickers, zero weights, or negative weights are dropped with a warning.
6. Duplicate tickers are aggregated by summing their weights before normalisation.
7. All remaining weights are normalised to sum to exactly 100%.
8. Constituents are ranked by descending weight and cumulative weight is computed.

Files missing required columns (ticker, name, sector, or weight) are rejected entirely. The engine will show an error message and will not proceed with partial data.

5.5 Historical Data Upload

The current upload flow is designed for a single current constituent snapshot. When a single snapshot is uploaded, the Trend Analysis page shows only that snapshot — there is no simulated drift to chart.

For a full 24-month historical trend using licensed data, your data provider should supply month-end constituent snapshots in long-form format with the same four required columns plus a date column:

```
date,ticker,name,sector,weight
2026-04-30,NPN,Naspers Ltd,Technology,10.7
2026-04-30,SBK,Standard Bank Group,Financials,5.4
2026-05-31,NPN,Naspers Ltd,Technology,10.5
2026-05-31,SBK,Standard Bank Group,Financials,5.6
```

The metrics layer already accepts long-form snapshot history directly — the compute_trend() function works on any long-form DataFrame with the correct columns. This means a historical upload can be wired in without changes to the analytics engine.

Ask your JSE data provider, custodian, or broker for month-end constituent weight files in CSV format. These are standard outputs from most institutional data platforms.

5.6 What Changes When You Upload

When a constituent file is uploaded and accepted, the following changes throughout the engine:

- The benchmark snapshot used by Dashboard, Holdings, Shock Scenarios, Portfolio Comparison, and Export Report all update to reflect the uploaded data
- The sidebar shows the as-of date from the uploaded file instead of the simulated data date
- The simulated data disclaimer watermark is replaced by the as-of date
- All metrics, charts, and risk scores are recomputed from the uploaded weights
- The Excel report cover sheet reflects the uploaded data as-of date

The mode automatically reverts to simulated data if the uploaded file is removed or if the session is refreshed without re-uploading.

6. Page-by-Page Tutorial

The engine has six pages accessible from the left-hand sidebar. In Simulated Data mode, use the Benchmark Index selector to switch between ALSI and SWIX. In Upload JSE Data mode, the benchmark is determined by the uploaded file.

6.1 Dashboard

The Dashboard gives you a complete picture of the current concentration profile in a single scrollable view.

KPI Cards

Six metric cards at the top: Risk Score, Risk Label, HHI, Effective N, Top-5 Weight, and Top-10 Weight. These six numbers summarise the entire concentration profile at a glance.

Composition Treemap

Constituents displayed as coloured rectangles grouped by sector. Rectangle size represents index weight. Hover to see the stock name and exact weight. When one or two rectangles visually dominate the chart, that is concentration risk made visible.

Cumulative Weight Curve

Plots cumulative weight against rank from largest to smallest stock. A steep curve crossing 50% before the sixth stock indicates severe concentration. Dashed lines mark the 50% and 80% thresholds.

Sector Distribution and Top-10 Bar Chart

Sector bar chart shows weight by sector. Top-10 bar chart shows individual weights for the ten largest constituents.

Concentration Risk Gauge

The semicircular gauge displays the Risk Score with colour bands for the four classifications. HHI and Effective N are displayed below as large numeric values.

> *In Upload JSE Data mode, the as-of date from your file appears in the sidebar and header instead of the simulated data disclaimer. This confirms the engine is running against your licensed data.*

6.2 Holdings

Full 40-name (or however many constituents are in your uploaded file) ranked constituent table. Columns: rank, ticker, full name, sector, individual weight, cumulative weight. Sortable by any column.

Cumulative weight colour coding: green below 50%, amber between 50% and 80%, red above 80%. Any row turning red marks the stock that pushed the index past 80% cumulative weight — a key risk threshold.

6.3 Trend Analysis

Shows 24 months of historical concentration data in simulated mode. In upload mode with a single snapshot, shows only the as-of date snapshot. Select any metric from the dropdown: Top-5 Weight, Top-10 Weight, HHI, Effective N, Risk Score.

The chart uses a dual Y-axis when HHI and a percentage metric are displayed simultaneously. Below the chart, a monthly snapshot table shows all metrics colour-coded by risk classification.

> *To get a full historical trend in upload mode, ask your data provider for month-end constituent weight files and upload the long-form historical CSV. The engine processes it directly without any code changes.*

6.4 Shock Scenarios

Pre-Built Scenarios

Six pre-built JSE stress scenarios: Naspers De-rating, Resources Rout, Big-Bank Stress, Dual Shock (Tech plus Resources), Gold Rally, Rand Depreciation Proxy. Each shows the estimated index impact in percentage points and how the Risk Score changes post-shock.

In Upload JSE Data mode, shock impacts are calculated from the uploaded constituent weights. A Naspers shock on a SWIX-weighted file will show a smaller impact than on an ALSI file because NPN has a smaller SWIX weight.

Custom Shock Builder

Select any constituent and apply an individual price move from -100% to +200%. The engine adjusts shocked names, renormalises all weights to 100, and recomputes all concentration metrics.

> *Sales tip: upload your client's actual benchmark data and run the Naspers De-rating scenario. Showing the exact rand-per-unit impact based on their actual portfolio benchmark is far more compelling than a simulated number.*

6.5 Portfolio Comparison

Demo Portfolio

With no CSV uploaded, the engine loads a demo portfolio from the top 15 benchmark names at equal weight (6.67% each). This is a deliberately less-concentrated portfolio that makes the comparison instructive.

Uploading a Portfolio

Prepare a CSV with two columns: ticker and weight. Weights do not need to sum to 100. The engine:

- Normalises column names to lowercase
- Uppercases tickers and strips whitespace
- Normalises weights to sum to 100
- Aggregates duplicate tickers
- Drops blank tickers, non-numeric weights, zero weights
- Rejects negative weights
- Retains tickers not found in the benchmark with Unknown sector metadata and a warning

Example portfolio CSV:

```
ticker,weight
NPN,25
```

```
SBK,15
BHP,10
```

Reading the Comparison

- Side-by-side KPI comparison — portfolio versus benchmark for all six core metrics
- Concentration gap table — delta between portfolio and benchmark (green = less concentrated, red = more concentrated)
- Grouped bar chart — portfolio weight versus benchmark weight for the top 15 names
- Sector weight comparison — overweight and underweight sectors versus the benchmark

When Upload JSE Data mode is active, the benchmark side of the comparison uses your licensed data rather than the simulation — making the portfolio analysis directly comparable to your actual mandate benchmark.

6.6 Export Report

Click Generate Report to build the workbook, then download it. Generated in memory — no file written to disk. The five sheets:

Sheet	Contents
Cover	Report metadata, key risk summary, risk scale legend, and as-of date.
Holdings	Full ranked constituent table with weight and sector.
Concentration	Full metrics table plus sector HHI breakdown.
Historical Trend	24-month monthly concentration metrics and risk labels.
Shock Scenarios	All six pre-built scenario impact tables.

In Upload JSE Data mode, the Cover sheet shows the as-of date from your file rather than the simulated data disclaimer. The Holdings sheet reflects your uploaded constituent data.

7. Sales Guide — Talking Points by Audience

7.1 Asset Manager CEO

Open with: The JSE is one of the most concentrated equity markets in the world. Your clients may not know that. We built a system to make it visible, measurable, and reportable.

- Benchmark-aware portfolios may be inheriting concentration risk without knowing it
- A single stock can account for 21% of the ALSI — one bad earnings release moves your whole fund
- Upload your own JSE data and the Risk Score reflects your actual benchmark on your reporting date
- The Excel report can go straight into your quarterly board pack without reformatting
- Built specifically for the JSE — not a global tool retrofitted to South Africa

7.2 Wealth Manager

Open with: When your client says they are invested in the ALSI, do they know they are essentially making a large bet on Naspers, Prosus, and Anglo American?

- Show clients what is really inside their equity allocation
- The Risk Score gives you a single number to explain portfolio risk in client-friendly language
- Upload the client's existing portfolio and show them side by side against the benchmark
- Use simulated mode for presentations; switch to uploaded JSE data for formal client reports
- Most advisers cannot produce this kind of analysis — this is a genuine differentiator

7.3 Pension Fund Trustee

Open with: Your equity manager says they are diversified. But if Naspers, Anglo, and FirstRand all move against you simultaneously, how much of your fund is at risk?

- Upload the official JSE constituent file from your data provider for reports that reflect actual index weights
- The HHI is the same measure used by competition authorities to assess market structure
- Use the Excel report in your investment committee minutes and trustee reports
- Supports your engagement with asset managers during quarterly review meetings

7.4 Institutional Investor

Open with: How much of your JSE equity return last year was driven by Naspers alone?

- Portfolio comparison lets you see whether active equity managers are genuinely adding diversification or replicating benchmark concentration

- Upload your licensed JSE data to run analytics against your actual mandate benchmark
- Sector concentration analytics show whether your total equity allocation is structurally overweight Technology or Resources
- Shock simulations support ICAAP-style stress testing without a quant team

8. Data Model Reference

Each benchmark snapshot — whether simulated or uploaded — contains the following columns:

Column	Description
ticker	JSE-style ticker code. Uppercased and stripped of whitespace.
name	Full constituent name. Falls back to ticker if blank.
sector	Sector classification. Becomes Unknown if blank.
weight	Constituent weight on a percentage scale. All weights sum to 100.
rank	Rank by descending weight. Rank 1 is the largest constituent.
cumulative_weight	Running cumulative weight from rank 1 downward.
date	Present in historical snapshots only. Monthly date for the trend series. In upload mode, read from the date, as_of_date, or effective_date column.

9. Technical Setup

9.1 System Requirements

- Python 3.10 or newer
- pip (Python package manager)
- Approximately 200MB of disk space
- Any modern web browser (Chrome, Edge, Firefox)
- No internet connection required after installation

9.2 Installation

Standard installation:

9. Place the project folder on your machine
10. Open a terminal and navigate to the project folder
11. Run: pip install -r requirements.txt
12. Run: streamlit run app.py
13. The dashboard opens at http://localhost:8501

With a virtual environment (recommended for production):

14. Run: python -m venv .venv
15. Run: .venv\Scripts\activate (Windows) or source .venv/bin/activate (Mac/Linux)
16. Run: pip install -r requirements.txt
17. Run: streamlit run app.py

9.3 Project Structure

File	Purpose
app.py	Streamlit dashboard — all six pages, data mode selector, sidebar, charts, upload flows, scenario UI, and report trigger.
modules/simulate_jse.py	Deterministic simulated ALSI/SWIX snapshots and 24-month historical snapshots.
modules/metrics.py	Concentration ratios, HHI, Effective N, composite Risk Score, sector breakdown, trend calculations.
modules/shock_simulator.py	Single-name and multi-name shock calculations plus the six-scenario stress library.
modules/excel_report.py	OpenPyXL report builder returning an in-memory .xlsx file.
tests/test_portfolio_comparison.py	Smoke test including portfolio comparison logic, upload validation, negative weight rejection, and JSE constituent file validation.
requirements.txt	Python package dependencies.

README.md	Technical readme for developers and deployment teams.

9.4 Smoke Test

From the project root: python tests/test_portfolio_comparison.py

The test verifies 12 conditions:

18. Benchmark snapshot and metrics load
19. Demo portfolio normalises correctly
20. KPI comparison values are available
21. Ticker and sector chart data builds without error
22. Concentration gap table is shaped correctly
23. Portfolio Comparison navigation strings match exactly
24. Duplicate upload tickers are aggregated
25. Negative weights are rejected
26. Single-name portfolios are classified as Critical
27. Shock simulations reject all-unknown ticker inputs
28. Uploaded JSE constituent data is validated, aggregated, and normalised
29. Uploaded JSE files with missing required columns are rejected

10. Roadmap

- Connect the historical upload flow directly to the Trend Analysis chart for multi-period uploaded data
- Add configurable benchmark universes and constituent metadata sources
- Add persisted report output or scheduled report generation for deployment environments
- Portfolio shock layer — run shock scenarios against uploaded portfolio holdings rather than only against the benchmark
- Broader automated test coverage for metric edge cases, shock scenarios, and Excel workbook generation
- Streamlit Cloud deployment for a hosted demo URL

11. Glossary

Term	Definition
ALSI	FTSE/JSE All Share Index. The primary South African equity benchmark covering all listed shares.
SWIX	Shareholder Weighted Index. Down-weights dual-listed stocks to reflect shares held by SA residents. Materially less concentrated than the ALSI.
HHI	Herfindahl-Hirschman Index. Sum of squared constituent weights. Ranges from 0 (diversified) to 1 (monopoly).
Effective N	Effective Number of Constituents. Calculated as 1 / HHI. The number of equal-weight stocks that would produce the same HHI.
Concentration Risk Score	QuantSpace composite score from 0 to 100, combining Top-5 weight (40%), HHI (40%), and Effective N ratio (20%).
Simulated Data Mode	Engine uses hardcoded JSE-realistic weights from simulate_jse.py. For demos and workflow testing only.
Upload JSE Data Mode	Engine uses constituent weights from a CSV uploaded by the user from their licensed data subscription.
As-of Date	The snapshot date of the constituent data. Read from the date, as_of_date, or effective_date column in the uploaded file.
Cumulative Weight	Running total of weights from largest to smallest constituent. Shows how quickly index exposure concentrates.
Shock Scenario	A simulation where one or more stocks move by a specified percentage and the resulting index impact is calculated.
Reg 28	Regulation 28 of the Pension Funds Act. Sets asset allocation limits for retirement funds including concentration constraints.
Cap-Weighted Index	An index where each constituent's weight is proportional to its market capitalisation. Structurally amplifies concentration over time.
Long-form History	A CSV with one row per constituent per date. Used for the Trend Analysis page when uploaded historical data is available.

This document covers both simulated and licensed JSE data modes. Simulated data is not investment advice and not official JSE data.