Data Analytics in Internal Audit

Applying data analytics in the Internal Audit Function of ING using KNIME

Tjasse Biewenga
KNIME Data Talk, May 2021
ING is a global financial institution with a strong European base, offering retail and wholesale banking services to customers in over 40 countries. Bank’s Headquarters are in Amsterdam (The Netherlands).

- ING has more than 55,000 employees. We have more than 39 million customers, with 13.9 million considered primary customers.
- Top 10 of European banks

“ING has set the ambition to become a data-driven organization with analytics as a business model”
Our Audit World

- 400+ staff, 35 teams, 20+ locations
- Global / local, expert teams

- Number of Auditable Entities: 2,781
- Average cycle: 2,6 yrs
- Duration audit: 8 - 10 weeks

- Number of Audits 2021: 450+

In >60% of audits analytics will be applied
Audits 2021
• Many different topics
• Many audits per topic in different locations

Audit Analytics
• Different tests for different topics
• Same tests with different data sources

Our Analytics Approach
• Data Analytics indispensable part of every audit
• All staff proficient in applying Analytics in their audit
• Expert role for Advanced Analytics and Support
Data Analytics – KNIME - Internal Audit - ING

People

• Auditors with DA skills (perform DA with least dependency)

Audit Process

• Large variety of audit topics and data sources
• Relatively short timelines
• Automation of repetitive tasks

Tool Requirements

• Friendly interface => intuitive UI
• Automation made easy => no scripting

• Re-usability => sharing and collaboration
• Flexibility and integration => file types, databases, APIs
• Scalable => no limits for data volume

• Low license costs => expected >350 users
Data Analytics – KNIME - Internal Audit - ING

KNIME Analytics Platform

KNIME Server
Data Analytics – KNIME - Internal Audit - ING

Use Case
Use Case Process Mining for KNIME Server

Analyzing Event Logs:
- Transaction id’s (“Cases”)
- Process steps (“Events”)
- Attributes to that events (e.g. timestamps, resources)

Use cases primarily:
- Process Discovery

Tools:
- CAS Process Mining node in KNIME Server
Process Mining
The purpose of the node is to present information regarding the process flow of a process dataset. The main output of the node is an interactive view containing a dashboard with different kind of information and interaction. The input columns should be indicated by the user, thus the proper configuration is used. Data are prepared to present clear transition from one status to another, together with transition duration. Initiating and Ending statuses are also hardcoded, and included in the process.
Use Case Process Mining for KNIME Server

Data Analytics – KNIME - Internal Audit - ING

- Remove (and) on next status
- Filter on path freq
- Remove (and) on previous status
- Add previous status and previous date column and process step counter col
- Remove DELETE_ IT statuses
- Group by prevStatus nextStatus
- Path Frequency
- Rename aggregations
- Resort to javascript required
- Normalise rank value
- Min / Max 0.5 and 4 reflecting stroke-width in graph
- Create add column rank
- Round rank to 4 decimals
- Node 1399
- Javascript view to display the graph

---

- Java Snippet
- String Manipulation
- Range Slider
- Filter Definition
- Component Input
- Component Output
- Java Snippet
- Rule-based Row Filter
- GroupBy
- Column Rename
- Column Resorter
- Normalizer
- Math Formula
- Math Formula
- Column Filter
- Generic JavaScript View
- Javascript View
- Graph View
Data Analytics – KNIME - Internal Audit - ING

The End

Thank you!