BIG DATA IN BANKING
Banks have large volumes of data available across multiple domains.

In any large organisation including banks the number of available use cases is both exhilarating and daunting.

Budgets, technology and capacity will guide the number of projects that can be attempted in a given period.

IT scheduling is critical to getting work plugged in. Once plugged in business change management is essential for adoption.
Data

External data
Credit bureaus, loyalty scheme partners, data vendors etc..

Risk & finance data
Rich & well curated risk data stores are widely available (credit scoring, Basel capital & financial reporting)

Geospatial Data
Available if you look (ATM, POS, APP).

Know your customer
Base data is good but in its infancy. Contact information could be better curated, e.g. active updating

Digital data
Web and app data is highly valuable but requires investment in infrastructure

Transaction data
Rich for primary banked customers with lots of opportunity for data mining (e.g. merchant spend analysis)
FOCUS IS GOVERNANCE & CONTROL

06. External Reporting

05. Management Reporting

04. Operational Reporting

03. Application Support

02. Guided BI

01. Self Service BI

FOCUS IS AGILITY
Data Governance

Must read GDPR & BCBS239

Data Ownership
Business owner & technical curator

Master data management
How the data comes together

Lineage
Path of information from source to application

Data quality
Profiling, exception reporting

Access rights
Permissions, masking, deletion

Metadata
Purpose, timeliness etc.
IT Decisions

01 PERSONAL COMPUTER VS SERVER
Minimise under the desk computing by creating appropriate server environment

02 TRADITIONAL SERVER VS DISTRIBUTED PROCESSING
Technical skills needed to establish and use environment e.g. Hadoop

03 MY SERVER VS RENTED
Convenience vs security of cloud. Also consider bandwidth

04 OTHER CONSIDERATIONS
In memory computing
Serverless architecture
Cold vs warm storage
Real time processing
USE CASES

**Know My Customer**
Single view of the customer, customers like me, ideal customer, features & triggers

**Know My Organisation**
Segment level BI, issues & opportunities, what we really do (listening)

**Know My Market**
Market view, benchmarking, micro segmentation, scenario analysis

**Manage Risk**
Credit, fraud, operations, churn, compliance

**Make Money**
New to bank, cross & up sell, pricing, product design

**Improve Experience**
Reduce friction, vend better, listen better
Do the right thing the right way
Understand the Customer

- Consistent vs single view
  - Portfolio view
  - Merchant view
- Clustering
- Features & triggers
  - Feature library
  - Value scoring
  - Pay away analysis
- What the customer wants
Personalisation

The use of data to personalise customer interactions and improve experience

A customer has a need

✓ I need to send money
✓ I need a place to live
✓ I need to pay for my child’s education

The need drives behavior

✓ Borrow money
✓ Cut back on spending & save
✓ Past behavior may have been driven by what we sold

Our purpose is to generate value

✓ Better user experience
✓ Responsible funding of life goals
✓ Advice in setting life goals
Personalization

Good use case for in memory computing

**Decision Engine**
Global rules, business rules, lead store, scores, optimization algorithm

**Channel Integration**
Real time & consistent, content store, ability to interact, fulfilment

**Customer Data**
Product holding, profitability, recent touchpoints, contact info.

**Measurement**
Response modelling, who responds to what and why
Customer Value Optimization

Let's not forget that banks are businesses

- **Income per customer**: Cross and upsell, Downsize, customer mix
  - %

- **Number of customers**: New to company, digital marketing, NTU analysis, reduce friction
  - %

- **Margin**: Product design, cost to serve (robotics, self service), behavioral migration, elasticity, underwriting
  - %

- **Term**: Value - real & perceived.
  - %

Value
Pricing & Profitability Tools
Credit Loss Reserving

Globally there has been a move to expected credit loss provisioning for accounting purposes (IFRS 9).

This is a combination of general insurance techniques:
• Probability of default (frequency)
• Loss given default (severity)

And life/pension techniques:
• Survival modelling to derive lifetime expected loss

Because of the data volumes (millions of rows) this calculation can take material time at banks often measured in days as opposed to hours or minutes.

Using big data techniques we have been able to perform the assumption derivation in seconds on gigabytes of data. This allows management to explore more microsegments to better manage credit risk appetite.
Credit Loss Reserving

Frequency of default calculation on 70gigs of data in a few seconds without caching.

This allows for more scenarios to be generated for more micro segments to better manage risk appetite.

Also useful for IFRS 9 to run multiple scenarios on the fly to determine bucket 1 bucket 2 segmentation.

Technique can easily be extended to runoff triangle generation and claims investigations by micro segments.
Non-Traditional Data

Voice data
What was said, how it was said, when was it said, in what context was it said

Email data
Categorization, routing, sequencing

System log data
Demand forecasting, sequencing, route cause analysis

Image data
FICA, identification, emotion

Analytix Engine
Natural language processing

Neural net example

Call centre logs

Emails

Classification of contact messaging:

- Root cause analysis
- Optimised routing
- Demand planning
- Focused training on high demand requests
How Change Happens

According to Google

1. **INDIVIDUAL CONTRIBUTOR**
   - One person does the task

2. **DELEGATION**
   - Many people do the task
   - The task needs to be defined and standardised

3. **DIGITISATION**
   - We teach a computer to do the task
   - Automate mundane
   - It needs to be well codified, e.g. ATM
   - Not every task needs to be digitised

4. **ANALYTICS**
   - Machine learning
   - Continuous review
   - Human in the loop
Team

Data Engineer
- Data production

Production Engineers
- Join the bits together
- App building

Data Scientist
- Widgets, (modules) APIs
- Classification, scoring, forecasting

Business
- Strategy, trade offs,
- Channel execution
Path to success

**Quality Production & Measurement**
Clear measurement of success buys further investment. Stable production avoids embarrassment and allows you to build on early wins.

**Be Use Case Driven**
Short sprints prioritized on ease of implementation, alignment with strategy and expected value.

**Build Strategic Capability**
Data, systems, people & process Invest in architecture & governance.

**Focus on adoption**
Business sponsorship & channel/customer user experience.
The VALUE of data lies in its ability to inform decisions.

The POWER of an organisation lies in its ability to execute its decisions.
Questions
Pravin Burra
CEO, Analytix Engine

Over 25 years of analytics experience

Ex partner at big 4 audit firm

Previous executive head of customer analytics at leading retail bank

CONTACT

Email: pravin@analytixengine.com
Mobile: +27 82 336 1515
www.analytixengine.com