IMF Experience with Big Data
Overview of the presentation

- Background – why are we now focusing on big data?
- Change – how are we responding to the challenge?
- Examples – what big data projects?
The Background – How we did things in the past.....

- Traditionally the IMF has depended almost entirely on structured data
- Two kinds of data:
  - Operational data
  - Official data
- Recurring crises –
  - Common feature – data gaps
  - IMF handling huge amounts of structured data
- This structure has posed a constraint on our efforts to use big data
How are we addressing the issue?

Recently, the IMF has introduced a number of initiatives to take advantage of the potential with big data.

- **2016**
  - Staff Discussion Note: “Big Data: Potential, Challenges and Statistical Implications”
- **2018**
  - New data strategy
- **2019**
  - Establishment of a new Data Services division in the Statistics Department with responsibility for big data
  - Creation of a Data Scientist position within the IMF “Job family”
  - Establishment of a “Community of interest” within the institution
  - Innovation lab
  - New Infrastructure
Big Data in the Fund

- New questions and new indicators
- Bridging time lags in the availability of official statistics
- Supporting the timelier forecasting of existing indicators
- As an innovative data source in the production of official statistics
Supporting big data is a strategy priority in the Fund

**New Data Strategy**

Approved by the IMF Executive Board on March 9, 2018
The Strategy Sets Three Recommendations on Big Data

- Address Skill Gaps
- Support Big Data in Fund Operations
- Peer-Learning
New organizational structure

- Review of our data function
- Established a Data Services Division in Statistics with responsibility for big data.
- Review of job competencies
  - Data Scientist
  - Data Architect
  - Data Analyst/Officer
Big data “Community of Interest”

- Inter-departmental forum to share learning and results
- Focus on developing skills as much as highlighting research
- A community of “equals”
  - Works outside the IMF departmental structure
  - No direct management review
New Innovation Lab

- Previous approach
  - Ideas for change reviewed through departmental structures
  - High costs of failure

- New approach
  - Innovation “competitions”
  - Share ideas (TED-Talks, forums)
  - Funds available to “try things out”
  - It is “OK to fail”

- Practical impact – lots of big data ideas came through the innovation lab
  - 109 ideas submitted
  - 10 ideas selected
  - 6 ideas tested with a proof of concept
Some big data projects

- Lots of decentralized projects by individual researchers
- Limited resources available
- Innovation lab
  - Providing resources
  - Know-how
  - Forum for results

SWIFT
- Global flow of funds

Sentiment indicators (Google Trends)
- Nowcasting; High Frequency Indicators

Week at the beach
- Tourism index

Scanner data (Indonesia)
- CPI, retail sales
SWIFT

- Built SWIFT messages database
- Provides real transaction data on payments and trade between countries around the world.
- Downside – access to this data is very expensive
Sentiment indicators

- Google
  - Private API provided by Google for Social Science Research.
  - Used to nowcast various macro-indicators.
- Central bank announcements
- Financial Times – “Fear” indicators used in the vulnerability exercise
Tourism data

- Composite price index (basket of items typically consumed on a beach holiday) in collaboration with TripAdvisor
- Recent examples of applications to Fund policy work include
  - WHD REO 2016,
  - 2018 Grenada article IV report
  - 2017 Maldives article IV.
Developed an interactive database available to research staff
Indonesia Scanner data project

- Joint project with Statistics Netherlands and Bank Indonesia
- Part of our Technical Assistance
- Objectives:
  - Use scanner data to obtain high frequency estimates of the CPI
  - Obtain more robust high frequency retail sales estimates
- Project ongoing – initial results are promising, especially for the CPI
Concluding remarks

- The institution has recognized the need to develop capacity in the area of big data.
- This required some organization changes:
  - New data strategy
  - Dedicated division
  - New career stream
- Lots of projects were underway, but decentralized.
- Need to create space for innovation
- Last issue – infrastructure – we are starting to think about it now......