

Self Introduction

Antranig Khecho

Lead Software Engineer in EPAM Systems Inc.

Enterprise Applications for Healthcare and Financial Applications

Republic of Armenia

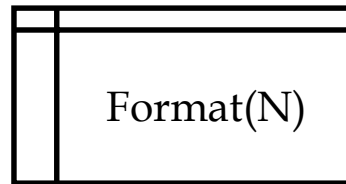
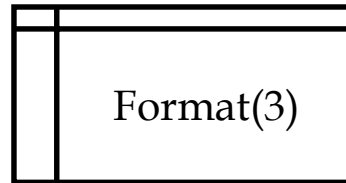
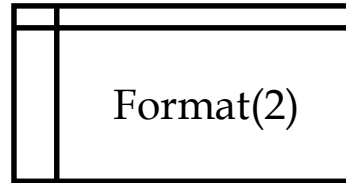
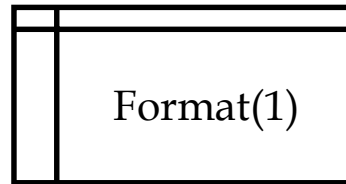
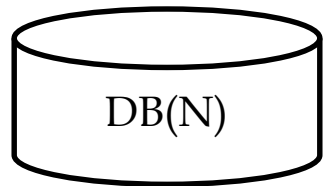
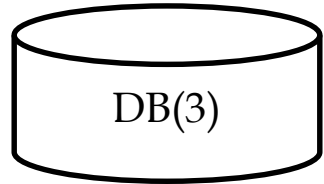
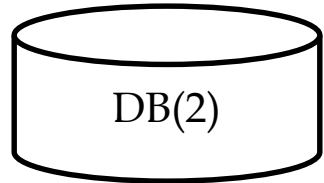
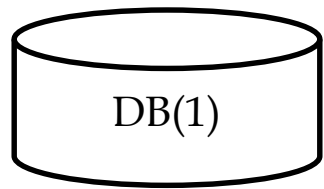
Agenda

1. Health Data Analysis Challenges
2. Health Data Analysis Proposed Solutions
3. Generic Industry Standard Platforms
4. Real Implemented Success Stories
5. Conclusion and Discussion

1. Health Data Analysis Challenges...

- *BIG*: Public health data are considerable and grow rapidly and exponentially. Consequently, accessibility and scalability challenges appears easily.
- *HETEROGENEOUS*: Public health data are either collected or generated from various sources or devices and stored in structured, unstructured, or even textual formats. As a result, it is a real bottleneck to integrate all these data together and define a unified manipulation approach for it.
- *HIGHLY CORRELATED*: One health data source can easily involves data that are with a high correlation and importance for data in other sources and those sources are strictly close for each other. So, how can find and approve this correlation.

1. Health Data Analysis Challenges.



2. Health Data Analysis Proposed Solutions.

- *Accessibility, Scalability, and Meaningful Integration:* We live in the age of *Big Data* and as we see there is no exception for health data. Consequently, we need to adopt a distributed system made up of a distributed storage infrastructure and offers a way to parallelize and execute analysis tasks on a cluster of machines.
- *Non-Trivial Correlation:* There is considerable amount of data analysis algorithms, approaches, and learning tools. We have to think about ones that try to analyze every possible correlation between each piece of data. In this context, it is a considerable to think about Association Mining.

3. *Generic Industry Standard Platforms.*

- *Hadoop Ecosystem.*
- *Elastic Search.*
- *Microservices Architecture.*
- *Microsoft Cognitive Toolkit, TensorFlow, etc. (for specific purposes).*
- *Association Mining Algorithms: Vertical Data Format.*

4. *Real Implemented Success Stories.*

- *Memorial Sloan Kettering Cancer Center: Historical data analysis platform to find probable rules for various cancer types causes.*
- *Stephanson Acquisto and Kolman Health Claims Data Analysis.*

Thank you for Listening!!