



Partners in Building UAE's Security & Economy

Exploring the Role of Modern
Identity Management Infrastructure
in Developing a Comprehensive

National Statistics

Presented by Dr. Ali M. Al-Khoury

In: Abu Dhabi Statistics Conference 2013
12-13 May 2013 | Ritz Carlton | Abu Dhabi | UAE

Federal Authority | هيئة اتحاديّة

Our Vision: To be a role model and reference point in proofing individual identity and build wealth informatics that guarantees innovative and sophisticated services for the benefit of UAE

www.emiratesid.ae © 2013 Emirates Identity Authority. All rights reserved



Agenda

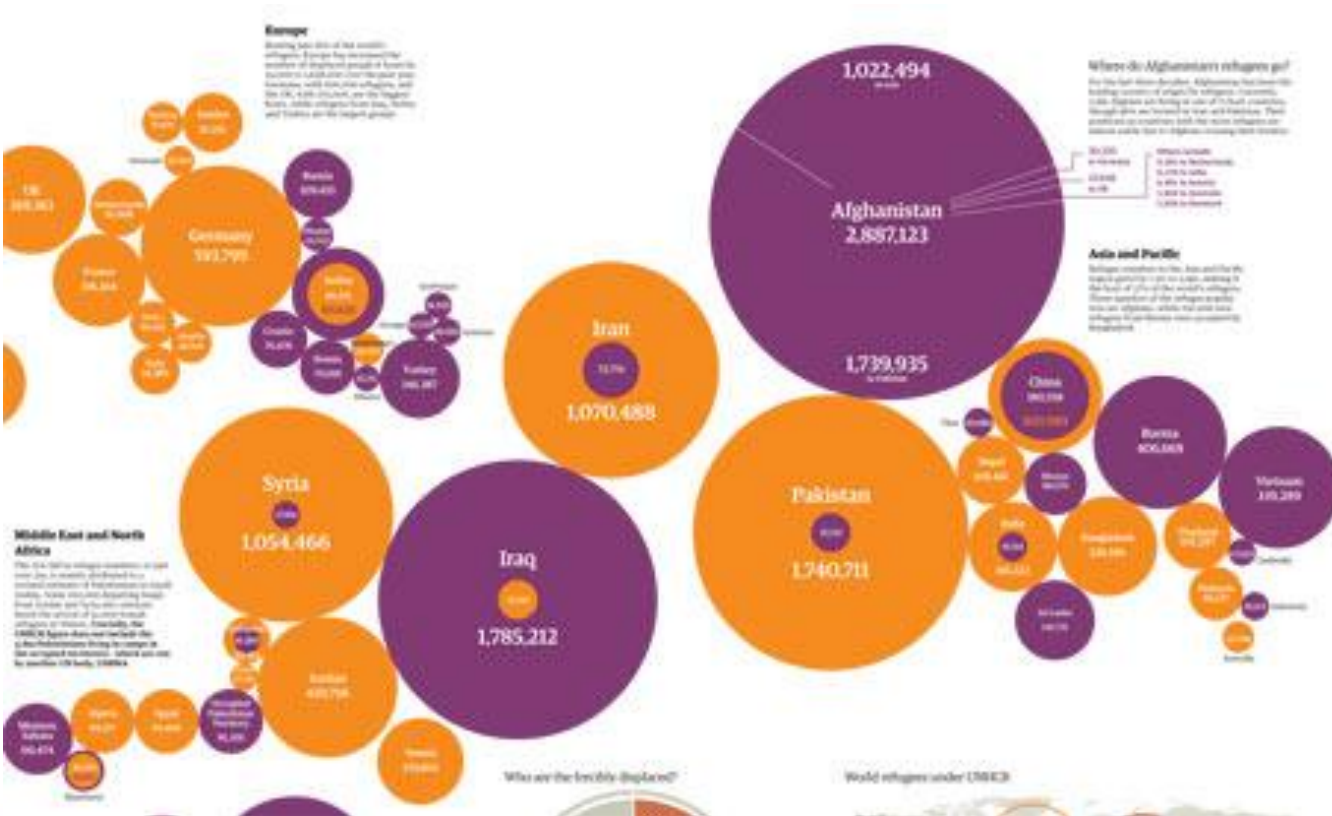
- Introduction: Statistics .. A Definition
- Statistics Scope and Boundaries
- Data and the Element of Uncertainty
- Modern Identity Management Systems
- Concluding Remarks

Agenda

- **Introduction: Statistics .. A Definition**
- Statistics Scope and Boundaries
- Data and the Element of Uncertainty
- Modern Identity Management Systems
- Concluding Remarks

Statistics: a classical definition...

A **mathematical subject** dealing with
“the **collection**, tabulation, **analysis**, interpretation
or **explanation**, and **presentation of data**”.....



Statistics: is a science ..

- A **science** in the realm of **mathematics**.
- Deal with **massive data**, **collate** into **meaningful** information
- **Historical** evidence
- Confident **predication** with a degree of **certainty**.



Information Age

- explosive growth in the **amount of data** created in the world, both structured and unstructured.
- It **continues to accelerate** and surprise us in terms of sheer volume.

Social media is producing a tremendous amount of new data.



The Age of Big Data ?



every day, we create

2.5 quintillion bytes of data ..

(2,500,000,000,000,000,000)

90% of the data in the world today has been created in the **last two years.**

IBM Report

We are truly in a

digital explosion era!

The Complexity of Context ..

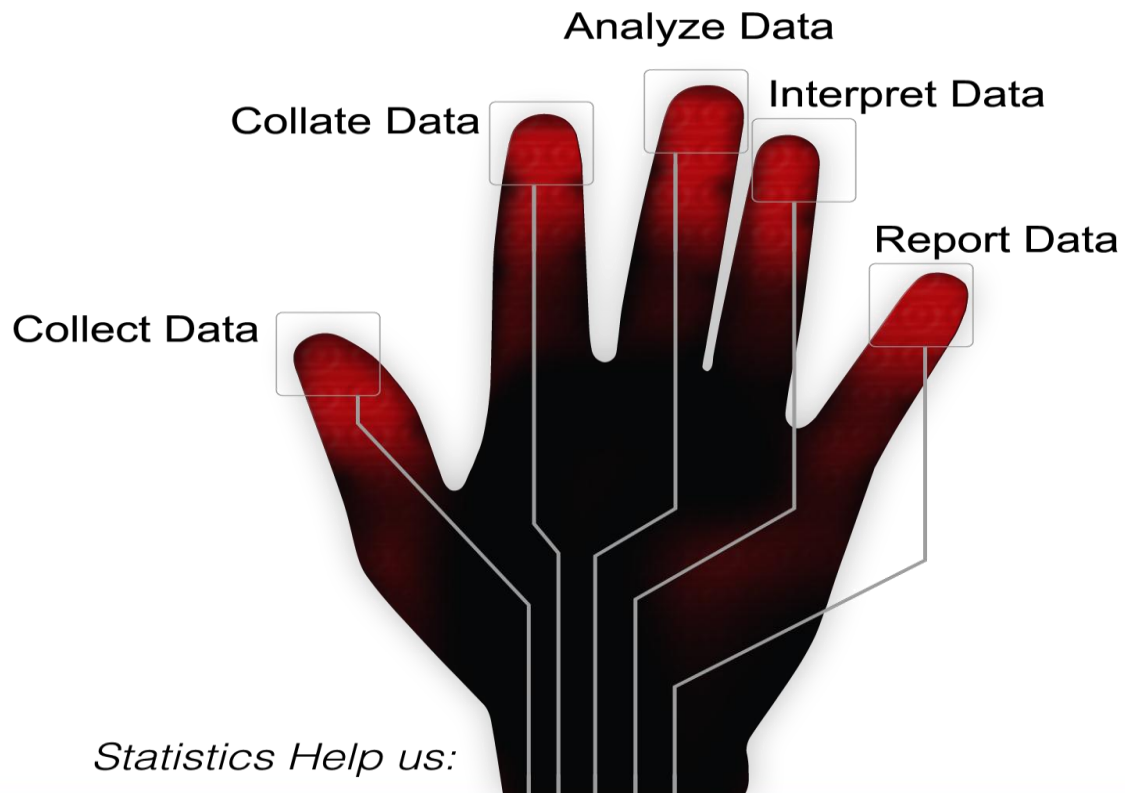
- Barrage of data **streamed at us!**
- Showering us with **massive rain of data?**
- Question: how can we create a **meaningful context?**
- Can we **filter this stream** into something more **structured** so that we can **make sense** of what we have and thus present it in a manner that is **useful?**



Agenda

- Introduction: Statistics .. A Definition
- **Statistics Scope and Boundaries**
- Data and the Element of Uncertainty
- Modern Identity Management Systems
- Concluding Remarks

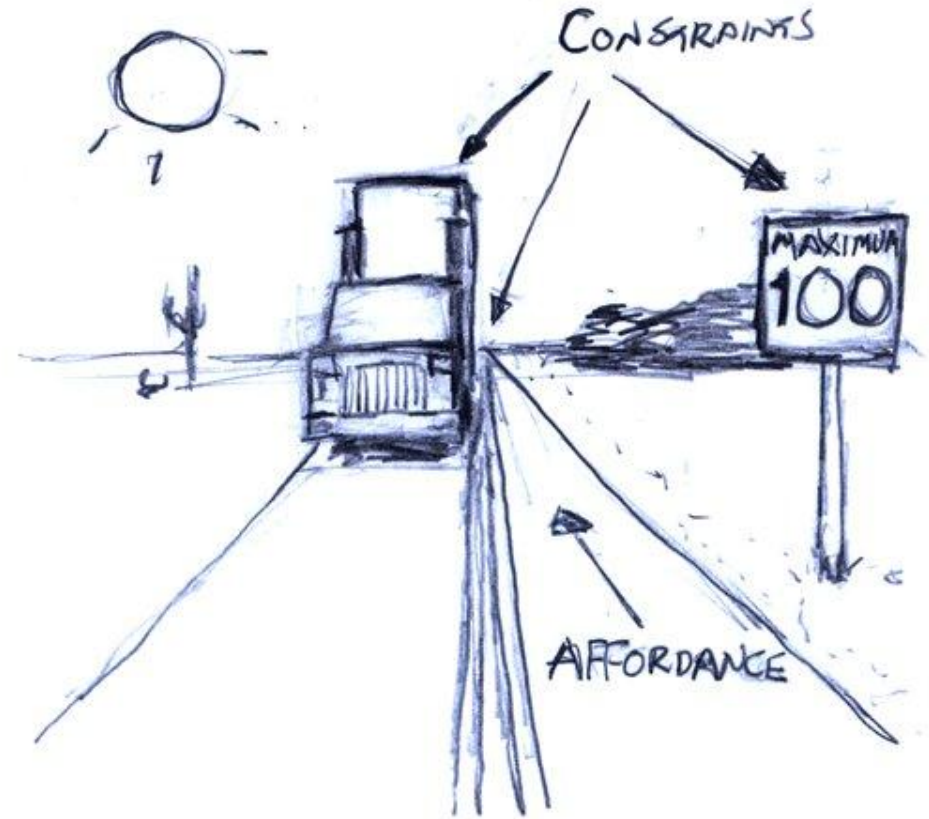
Statistics helps us in structuring the data into **matrices** and truly **brings us closer to a deeper understanding** that is **hidden** in the numbers associated with the data.



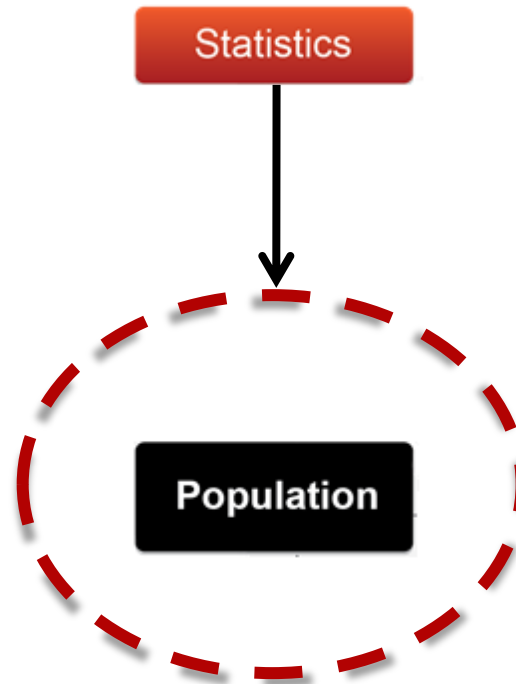
Statistics Help us:

enhance
understanding
& constitute
knowledge
(meaningful
information).

- **Boundary or constraints** under which a Statistical Analysis is reported



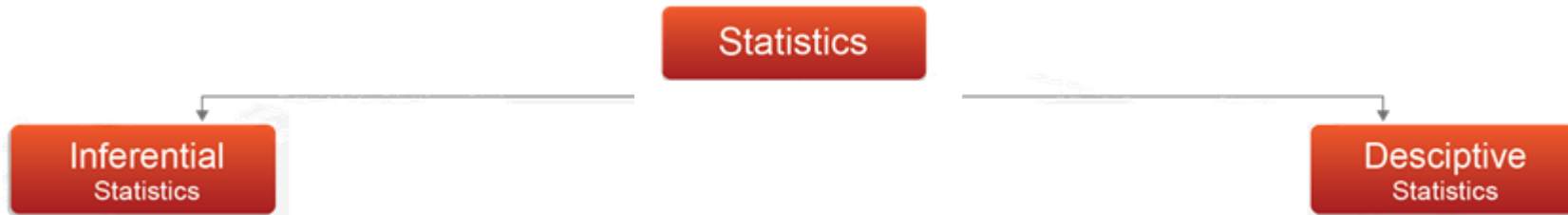
The People Dimension



Population in statistical terms is the **key subject of interest**, specifically in **demographic analysis**.

Scope of Statistics

Statistics spans over several distinct phases



Draw conclusions on unknown process parameters based on information.

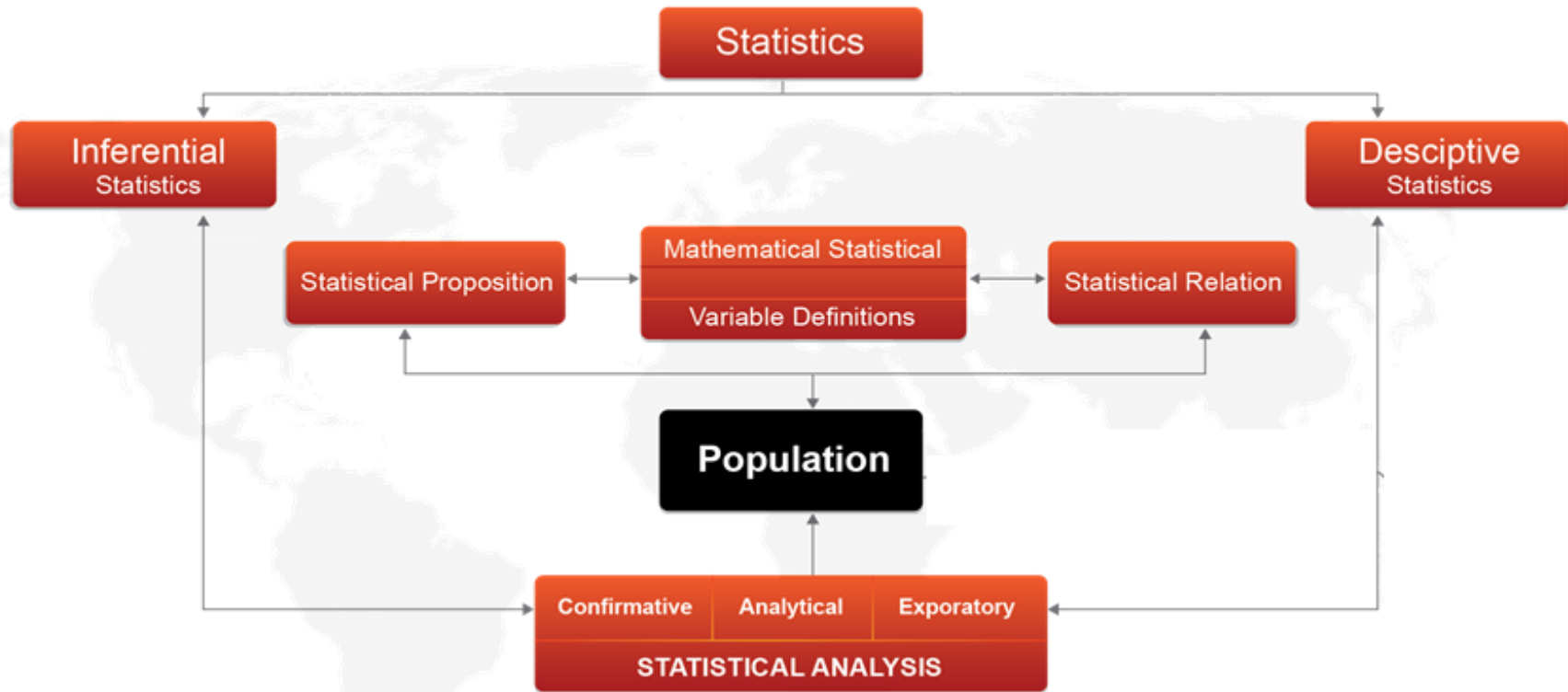
Uses probability, leading to either a **confirmation or rejection** of the hypothesis.

Population

Describes the **characteristics** of a **product or process** using information collected on it.

Hypothesis formation stage where “statistical thinking” plays a major role.

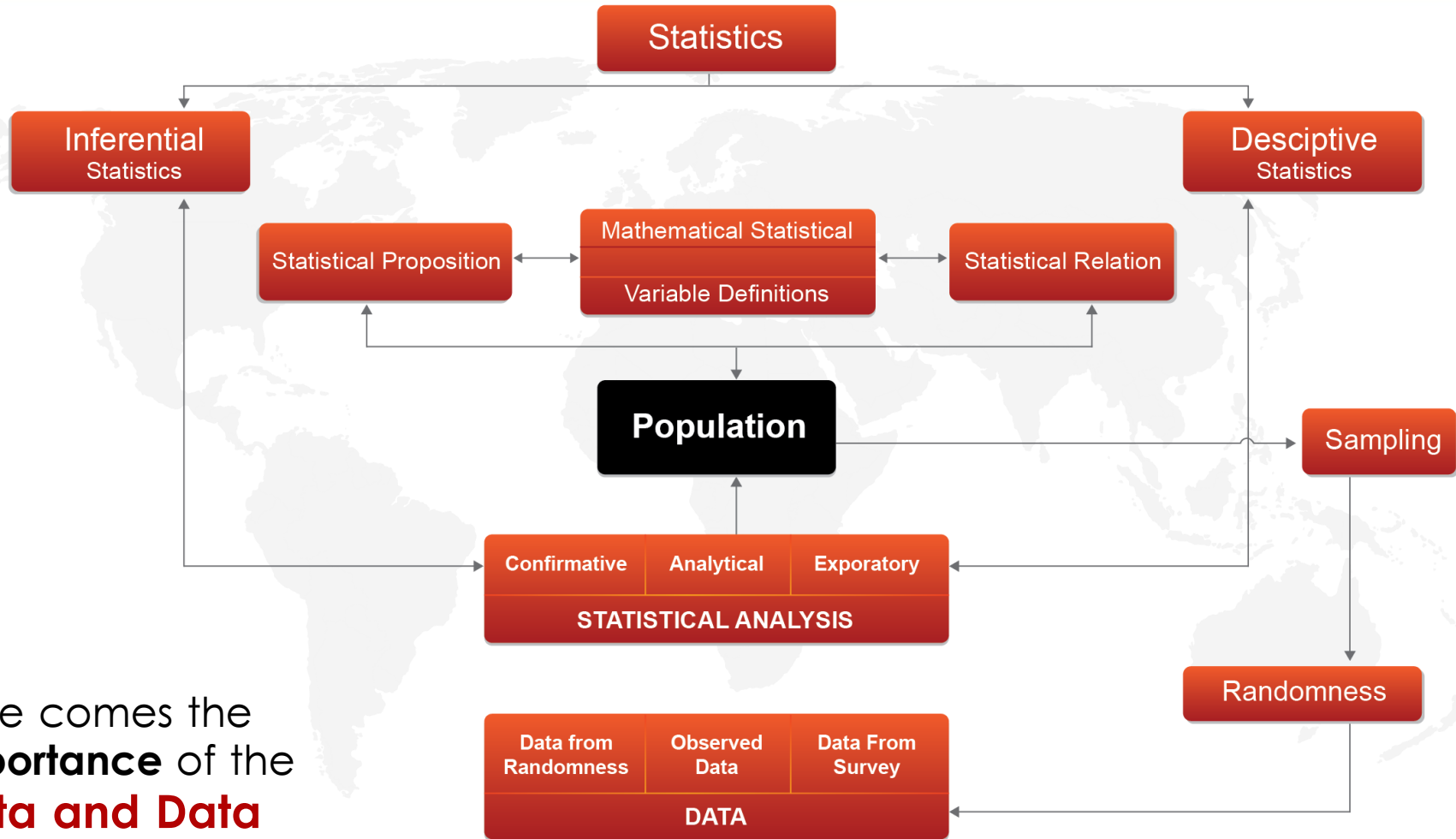
Scope of Statistics



Analytical with Validation of the Population

“Scale Up” or the **Extrapolation Analysis**.

Importance of Data (Collection)



Here comes the **importance** of the **Data and Data Collection**.

Depending on the size of the population, statisticians would draw a sample from the population for the required analysis.

Agenda

- Introduction: Statistics .. A Definition
- Statistics Scope and Boundaries
- **Data and the Element of Uncertainty**
- Modern Identity Management Systems
- Concluding Remarks

Data is only
as good as
the
collection
scheme

Data in statistics is an **abstraction of the population** and has three levels- raw.

Knowledge

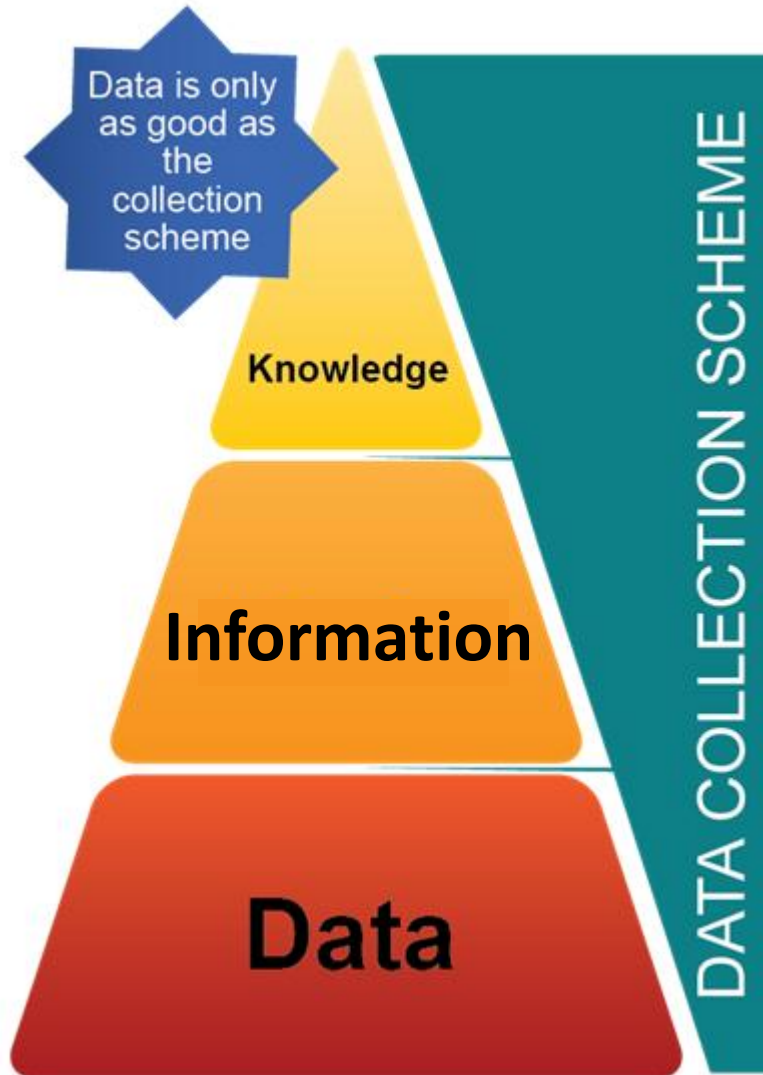
Knowledge at the highest level.

Information

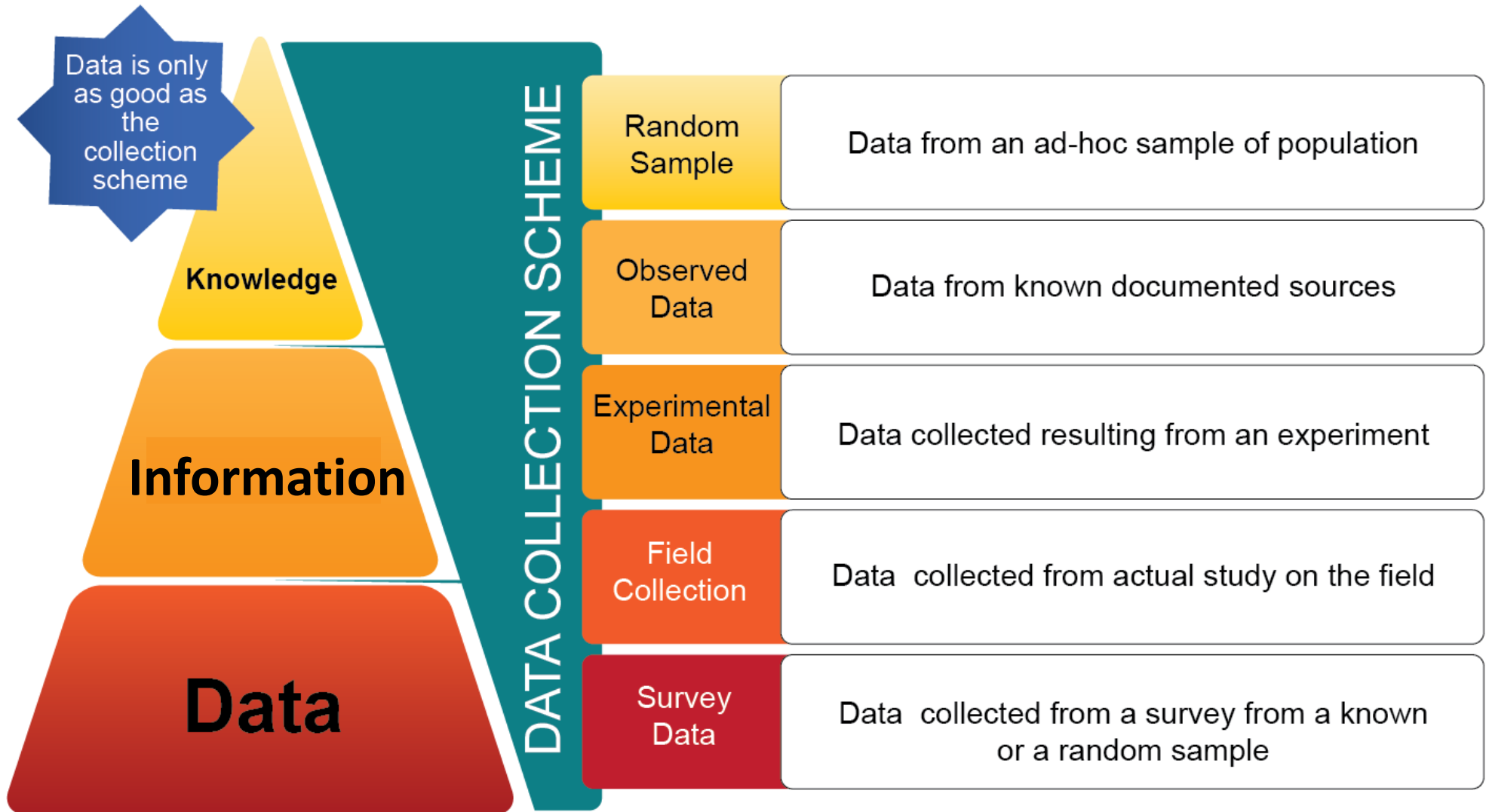
Information at the next level

Data

Data as the lowest level of abstraction

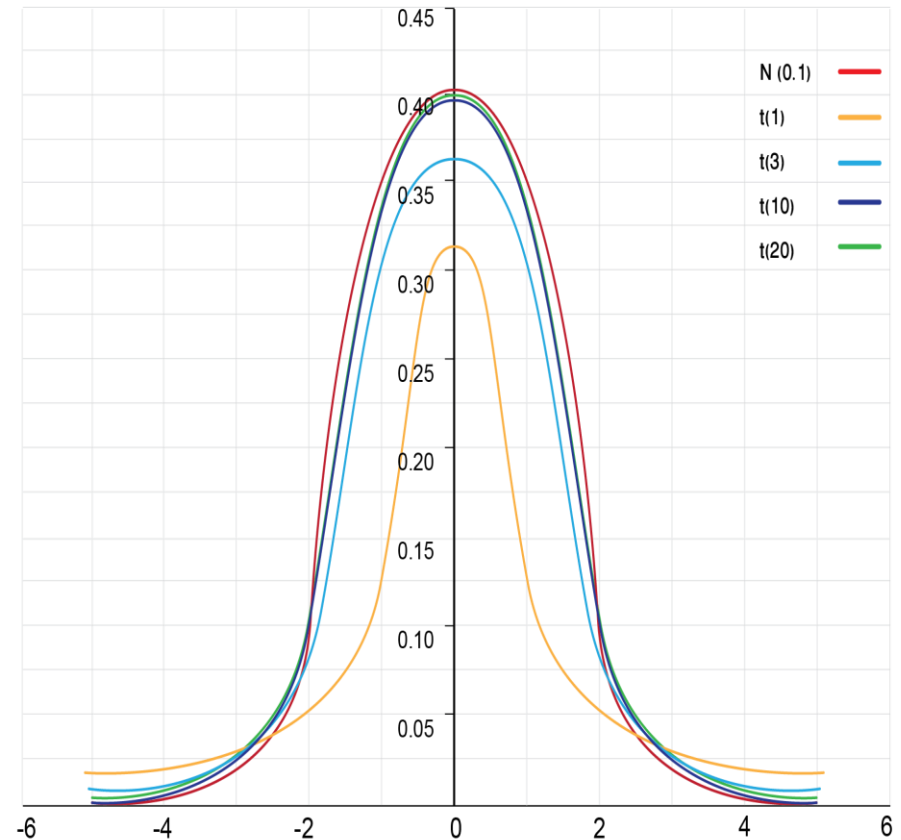


Data abstraction depends on the **scheme of Data Collection ..**



Uncertainty in Statistics

- Statistics by definition is an **approximation**.
- It is an **estimate**
- It is based on **probability**
- Utilizes **distribution analysis**
- Since we use hypothesis and sampling, there are always **variables** that depend on **assumptions**

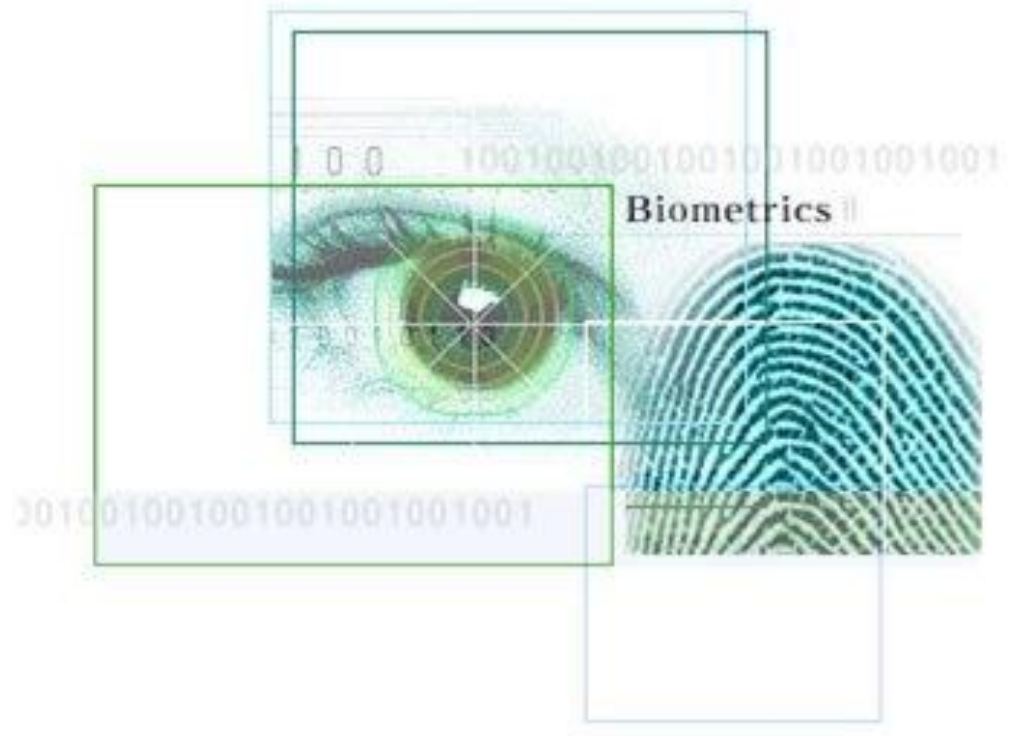


Agenda

- Introduction: Statistics .. A Definition
- Statistics Scope and Boundaries
- Data and the Element of Uncertainty
- **Modern Identity Management Systems**
- Concluding Remarks

Modern Identity Management Systems

- Many government launched modern **identity management** infrastructure.
- Use of advanced **identification and authentication** technologies.



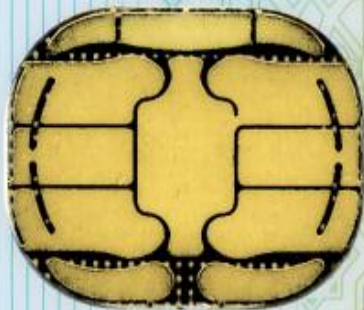
United Arab Emirates



دولة الإمارات العربية المتحدة

Identity Card

بطاقة هوية



ID Number / رقم الهوية

784-1968-6570305-0



الإسم: احمد محمد عبد الله

Name: Ahmed Mohamed Abdulla

الجنسية: الإمارات العربية المتحدة

Nationality: United Arab Emirates



Modern Identity Management Systems

**Formation &
Development**
2000–2005

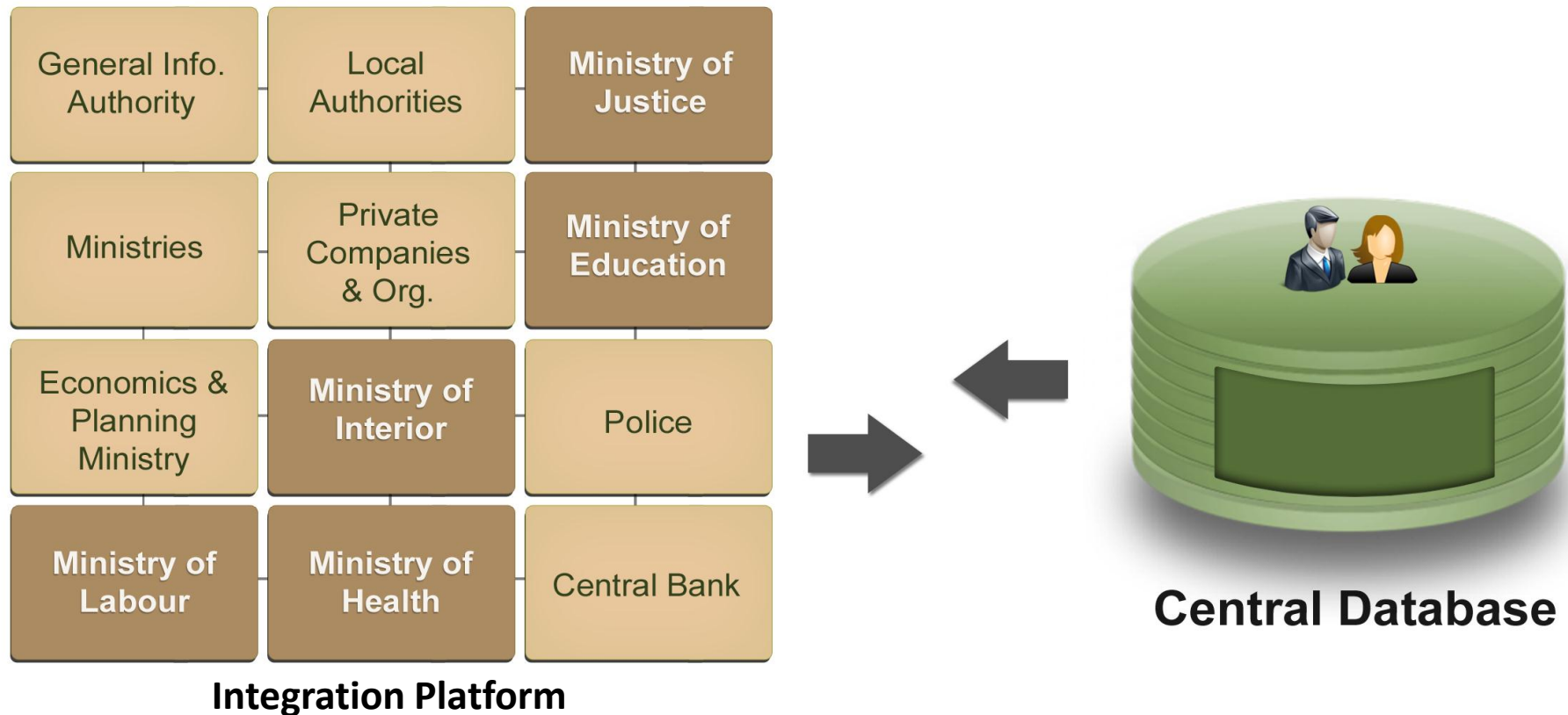
Enrolment

99%
Enrolled
Population

June 2005

Dec 2012

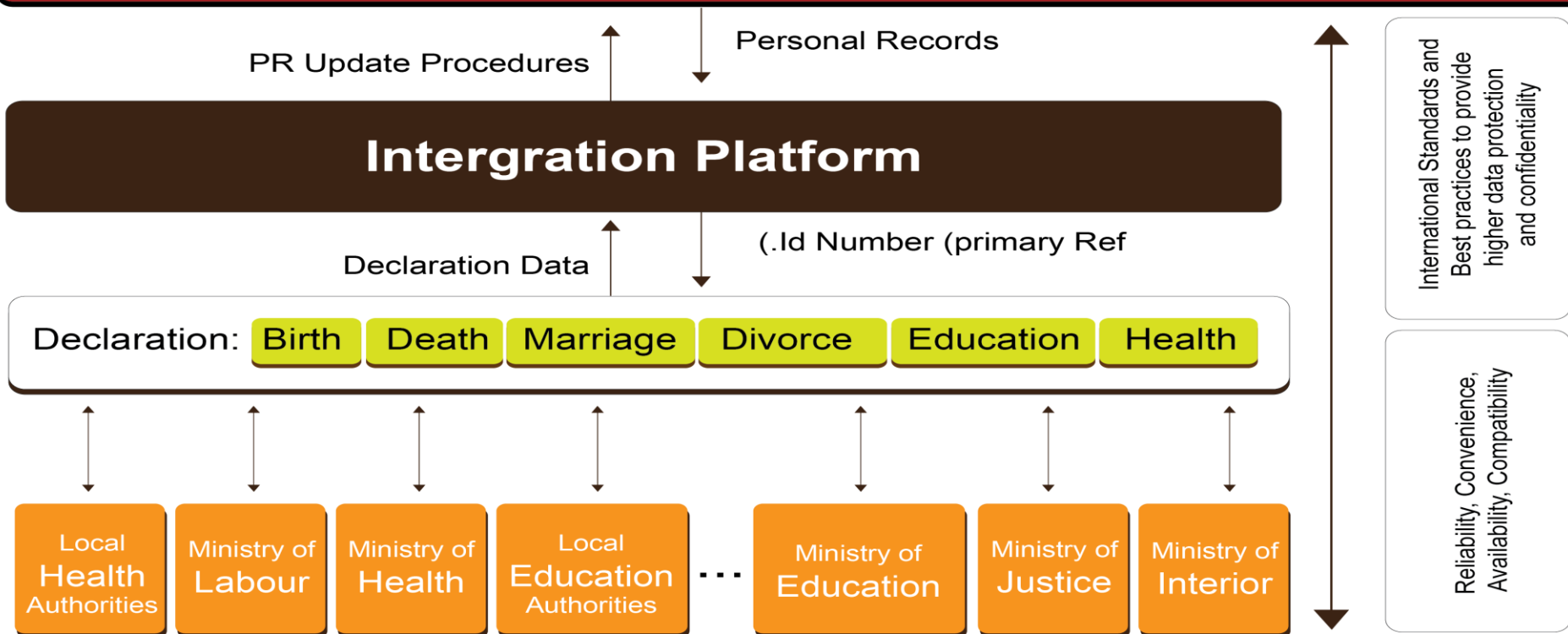
Keeping Population Register **up to date** **Support Decision Making**



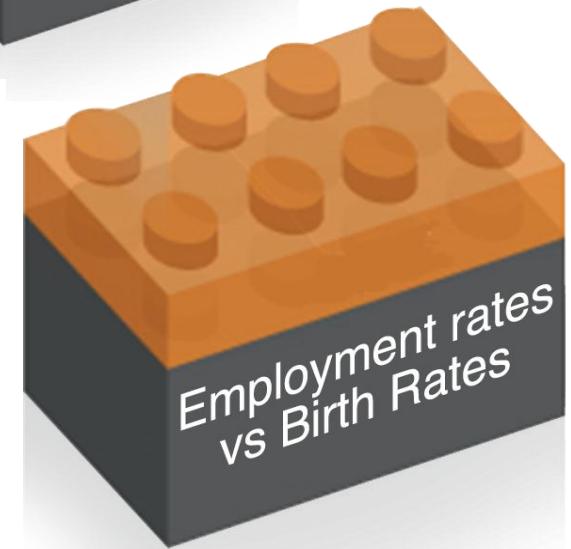
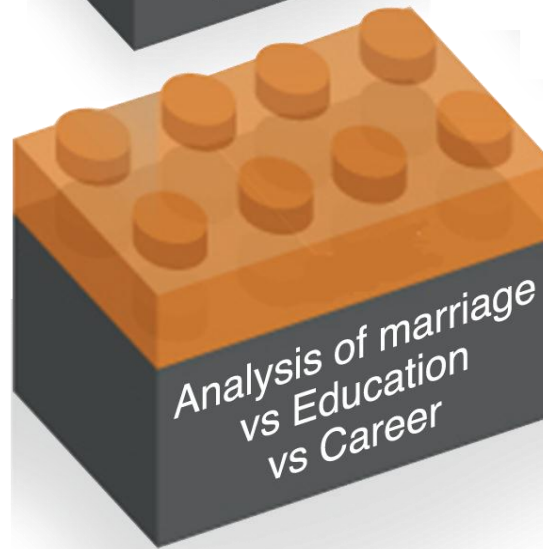
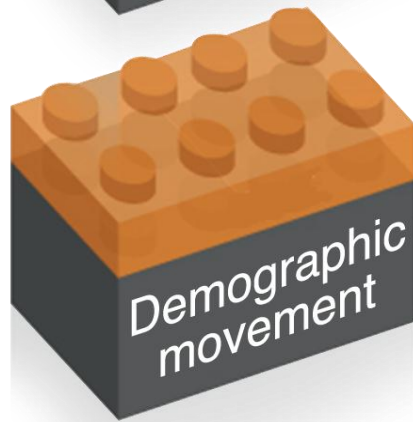
Integration Platform Architecture



Population Register System



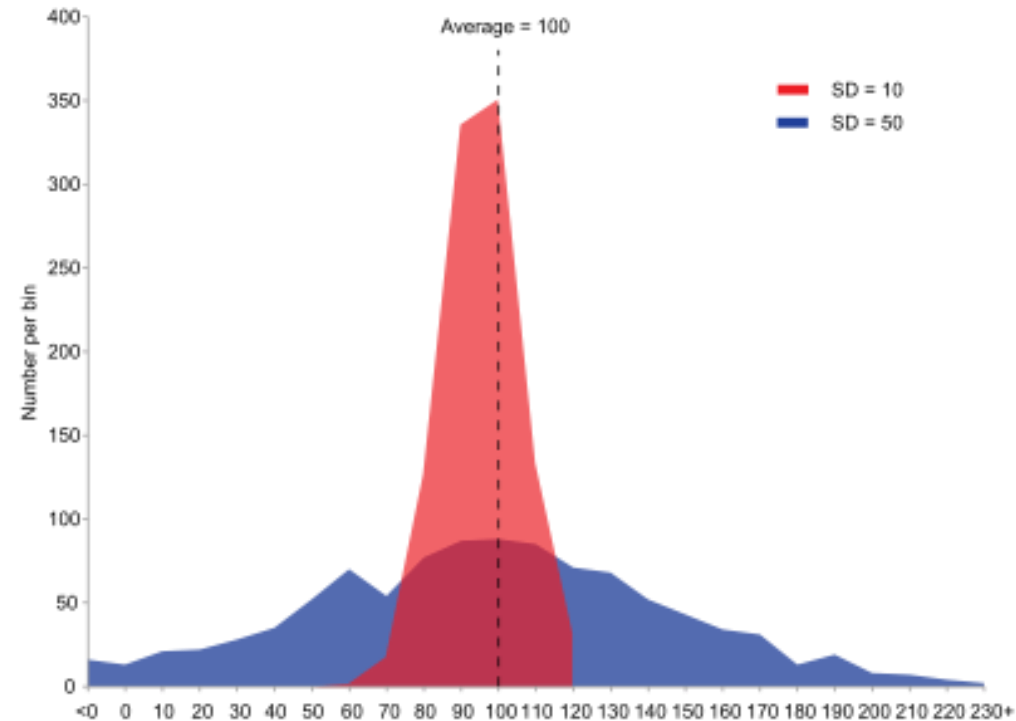
Examples of Statistical Analysis



(Un)Certainty in Statistics

The base of statistical analysis:

- **Confidence Interval:** provides the **measure of certainty**
- **Probability:** provides the **measure of uncertainty**



Statistical Analysis =

f{ Variables, Binomial relations, CI, P}

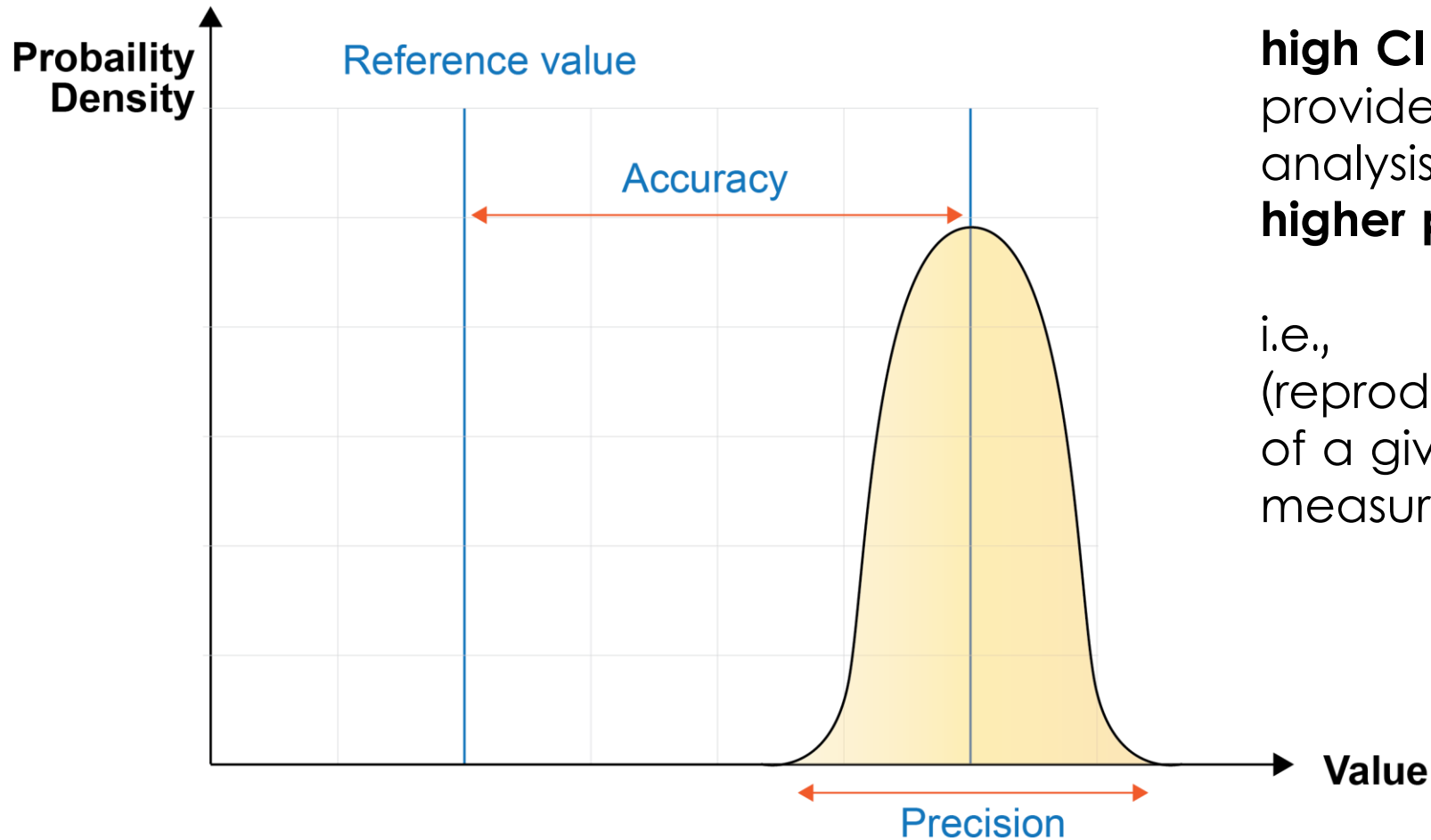
- Function of **considered variables** and **relationships**.
- Bound by the **Confidence Interval** (CI) and the **Probability** (P) of the occurrence of an event in the population.

Statistical Analysis =

f{ Variables, Binomial relations, CI, P}
limit $p \rightarrow 1$ f(Statistical Analysis) = CI

- There is a limit to how true is the analysis
- This limit is defined by the **Probability and the Confidence Interval.**
- **Probability could be 100% (1) but the true limit is the Confidence Interval (CI).**

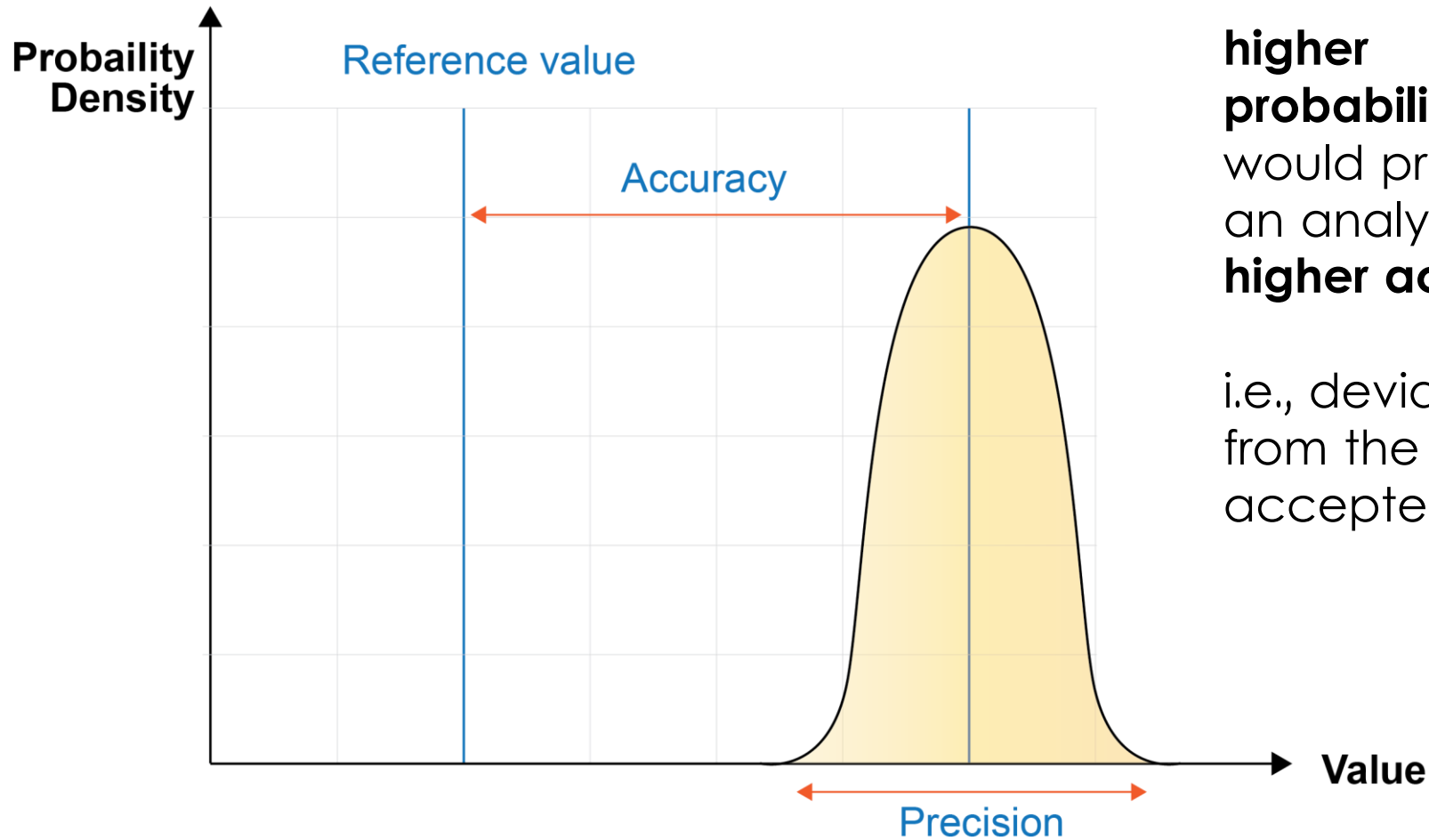
CI and P: Accuracy vs Precision of Statistical Analysis



high CI would provide an analysis with **higher precision.**

i.e.,
(reproducibility)
of a given
measurement

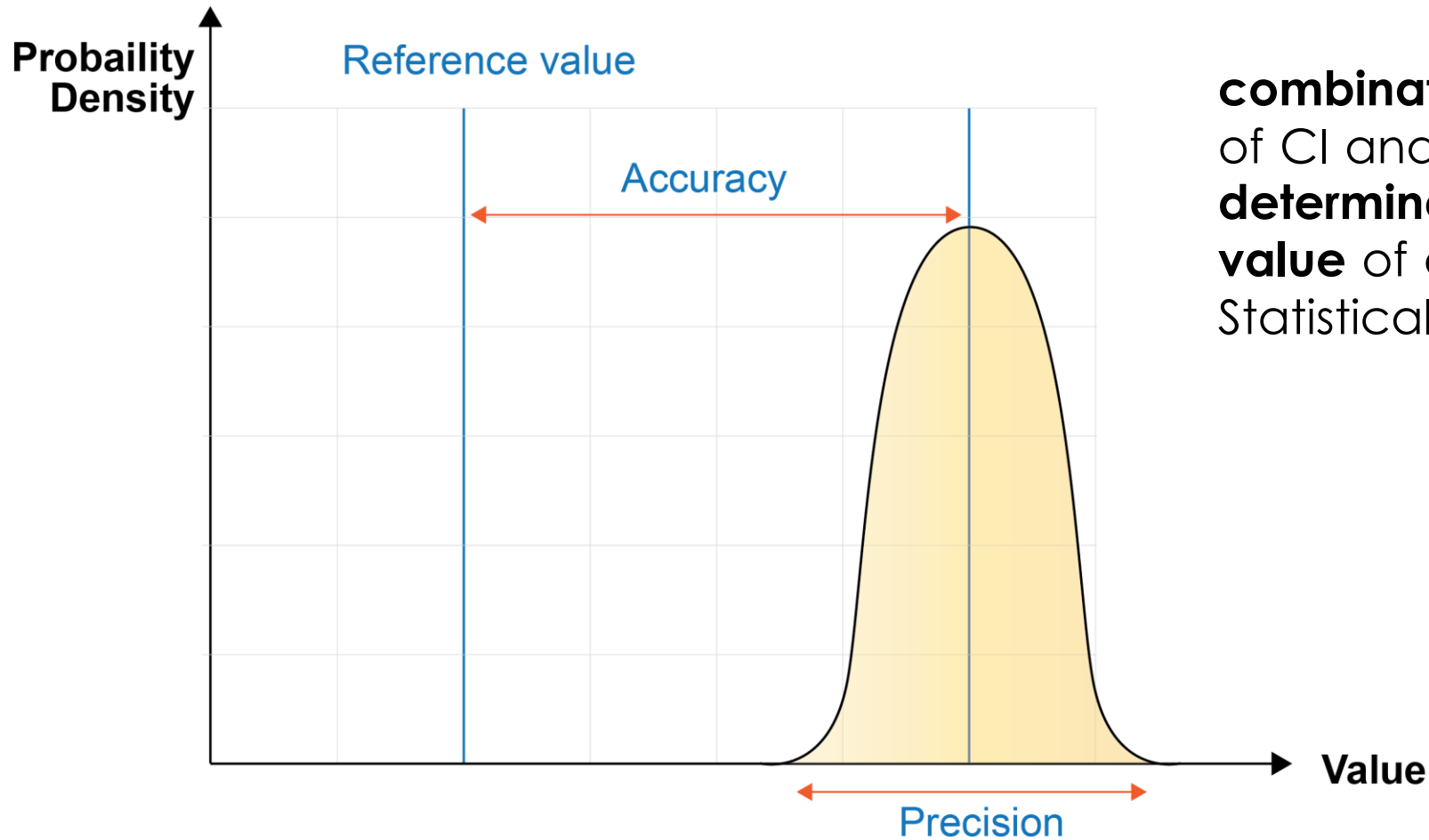
CI and P: Accuracy vs Precision of Statistical Analysis



higher probability
would provide
an analysis with
higher accuracy

i.e., deviation
from the
accepted value

CI and P: Accuracy vs Precision of Statistical Analysis



combination
of CI and P
determine the
value of a
Statistical Analysis.

Statistical Analysis = f{ Variables, Binomial relations, CI, P},

$$\text{limit}_{P \rightarrow 1} f(\text{Statistical Analysis}) = \text{CI}$$

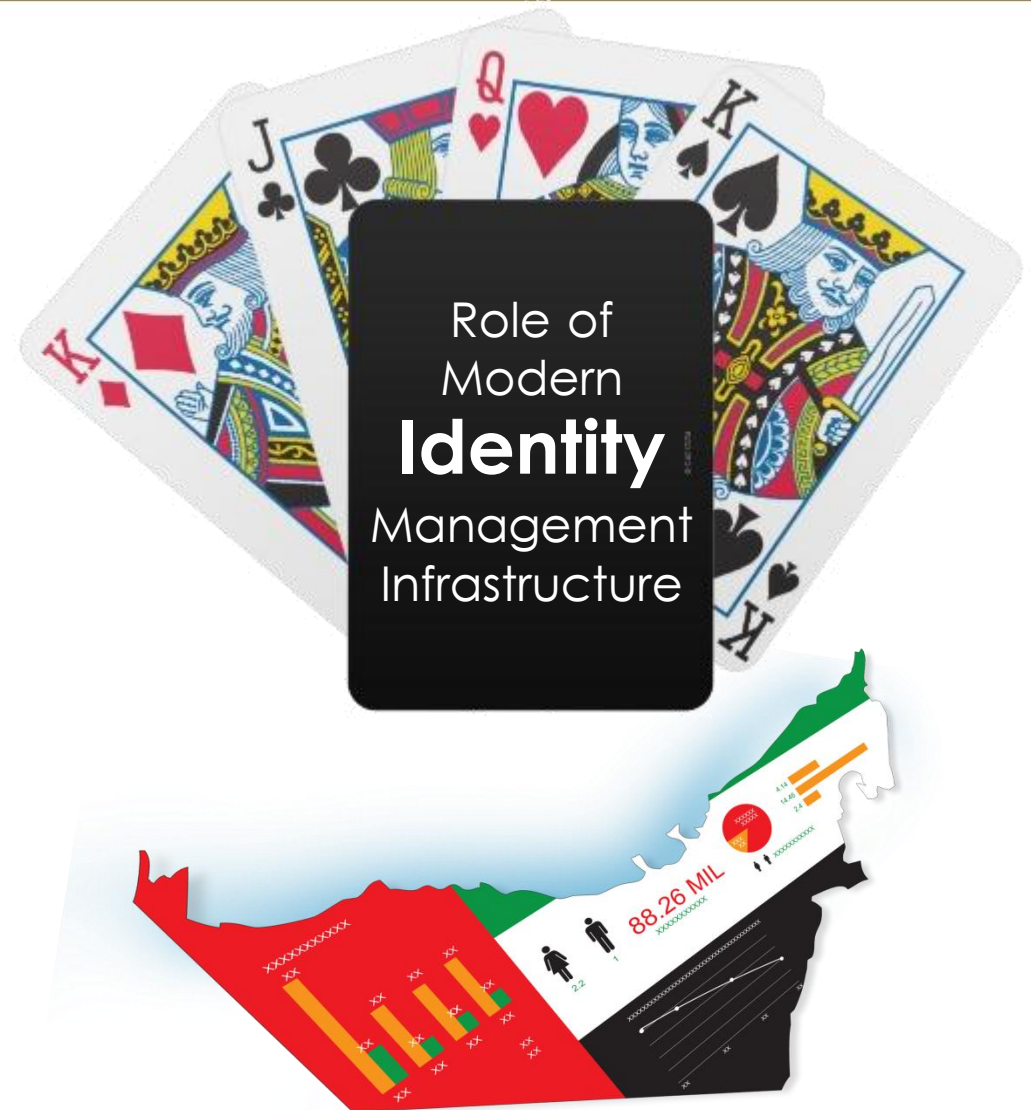
Corollary of the above

As CI $\rightarrow \infty$, f(SA) \uparrow pushing the CI higher

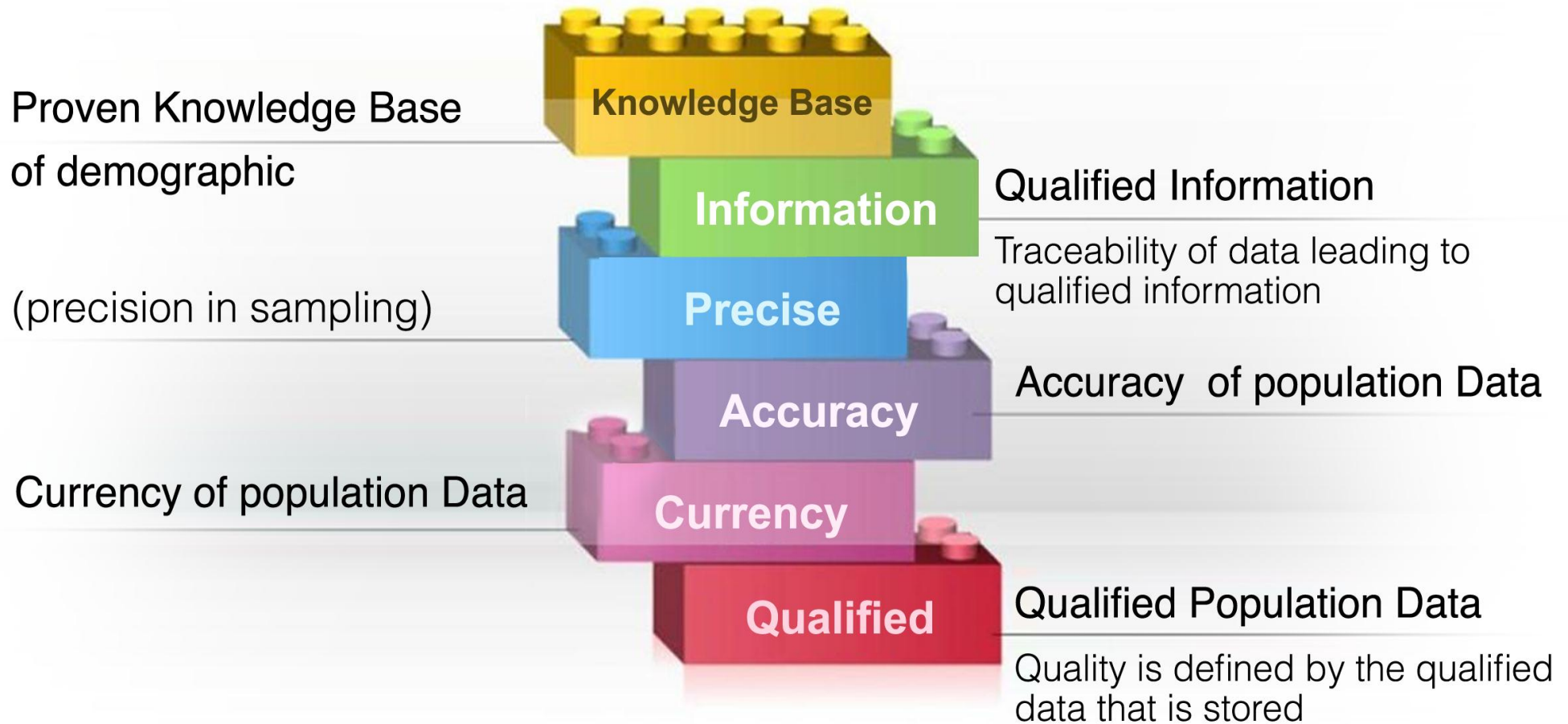
$$\text{since limit}_{P \rightarrow 1} f(\text{Statistical Analysis}) \rightarrow \infty$$

Role of Modern Identity Management Systems

- Increase the **Confidence Interval**
- Improve the **Probability**
- Enhance the **Analysis**
- Reduce **Uncertainty**
- Reinforce **Predictability**
- Exalted **Data Collection** from qualified data.



Quality of Data in Statistics



Agenda

- Introduction: Statistics .. A Definition
- Statistics Scope and Boundaries
- Data and the Element of Uncertainty
- Modern Identity Management Systems
- **Concluding Remarks**

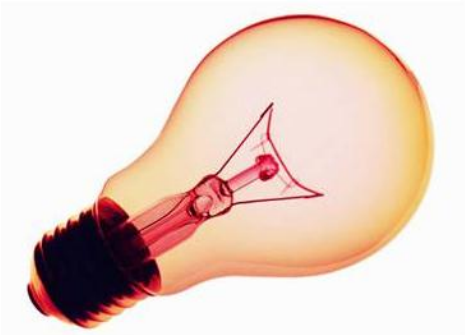
Concluding Remarks

- Statistics should be viewed to be about **building Knowledge** to support social, environmental, and economical development.
- i.e., to **improve efficiency, speed & accuracy of decisions, ability to forecast, cut costs, save energy, improve services, optimise infrastructure, enhance citizens quality of life, reduce environmental footprints, fuel innovation and diverse sustainable economic growth.**



\$300 billion

potential annual value to US
healthcare.



Challenges ...

- Many **national statistical systems** are not able to **effectively** meet data demand.
- **Constraints** include: inadequate **institutional & human capacities, infrastructure, funding**, etc.



By 2018,
United States will face
shortage of around
2 million
managers and analysts
with know-how to
create and use data to
make effective decisions.



Data Quality Challenges

- **International pressure** to form **higher quality** government statistical information.



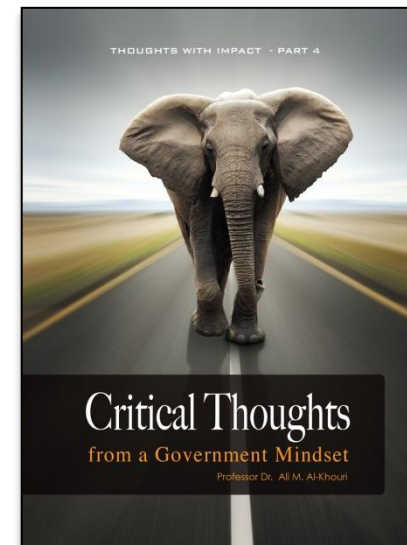
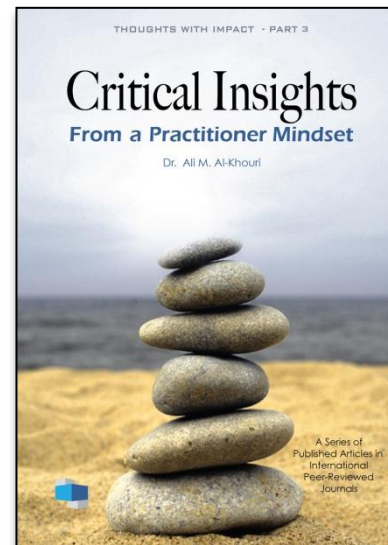
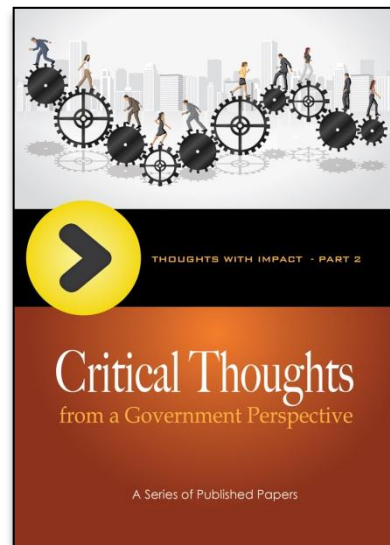
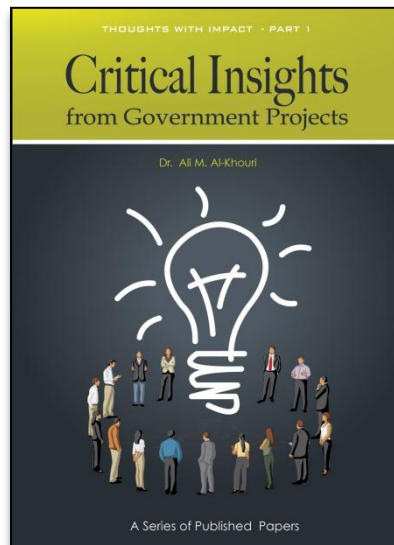
- Current strategies and approaches are likely to **change**.
- Governments will be forced to **re-examine** their **traditional systems** and consider new ways and approaches to **statistical analysis**, and use of **Big Data**.



More Information:

Read our recent research from:

<http://www.emiratesid.gov.ae/ar/media-center/publications.aspx>



“Data is a new class of
economic asset,
like currency and gold.”

World Economic Forum 2012

Thank You

Dr. Ail M. Al-Khouri
Director General
Emirates Identity Authority
United Arab Emirates
ali.alkhouri@emiratesid.ae
www.emiratesid.ae