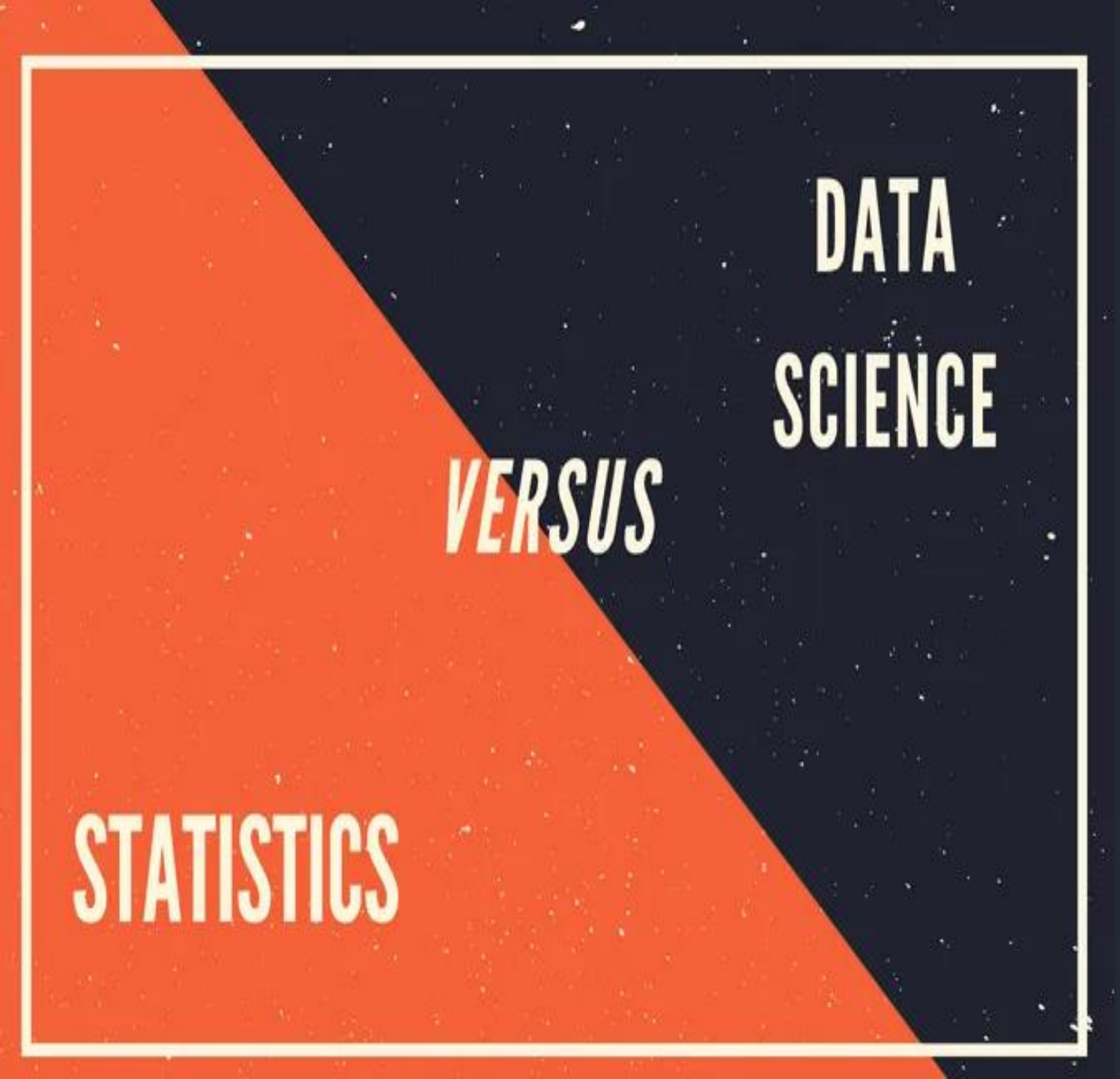


what’s the future?

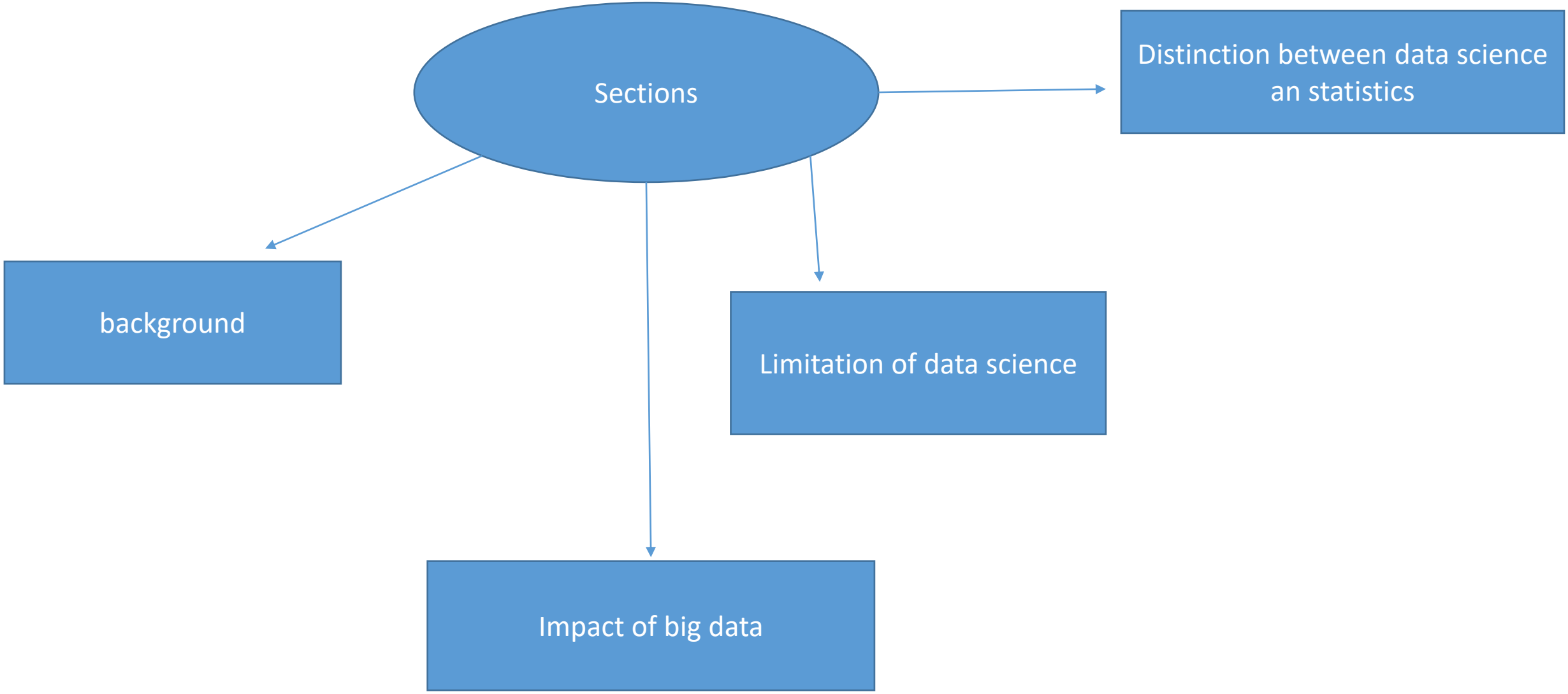
By hossein hassani

Student:salma akram
Email:A00022105@Sorbonne.ae

Hossein Hassani

Dr. Hossein Hassani has been included in a recently published list by Stanford University and Elsevier of the top 1% of world scientists across various academic disciplines. Currently, Dr. Hassani serves as an Adjunct Professor at Webster University in Vienna, Austria.

The article's section



Background: statistics and data science

- **Historical Development**
- **Big data's impact**

  Google's Chief Economist Hal Varian emphasises, the complexities inherent in modern world problems demands something more than statistics for understanding and extracting value

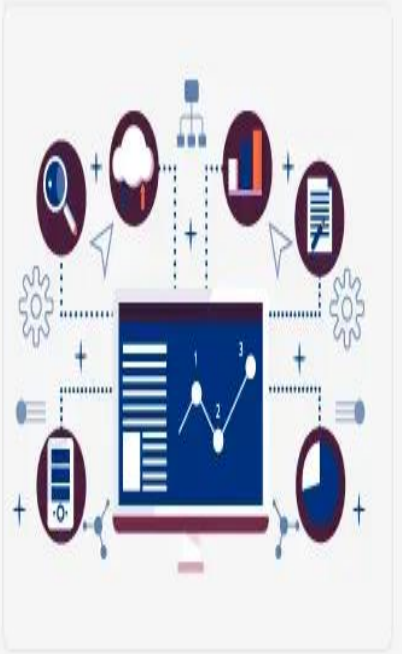
  from data

- **Importance of Statistical Skills**

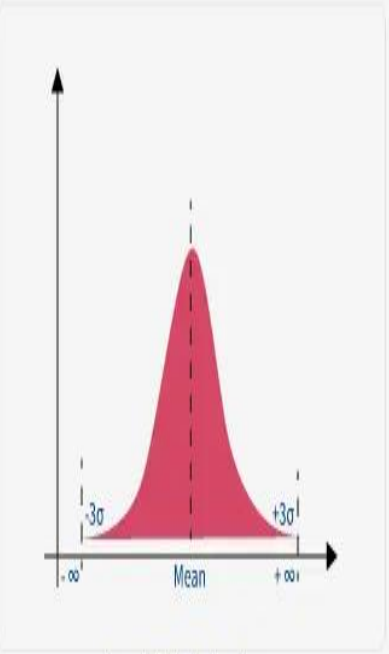
  the ability to use statistics to infer insights from smaller datasets onto larger populations and the importance of understanding Bayesian probabilities for machine learning.

- **Blurring boundaries**

  the emergence of data science has blurred the distinction between statisticians and non-statisticians in the field of data analysis.



DATA SCIENCE



Normal Distribution Curve

STATISTICS

- **Relationship Between Statistics and Data Science**
- **Debate Over the Relationship**

Some argue that data scientists downplay the importance of traditional statistics, while others see statistics as a fundamental component of data science.
Some believe that statistics is the foundation of data science, an it should be renamed data science while others argue that data science encompasses more than just statistics.

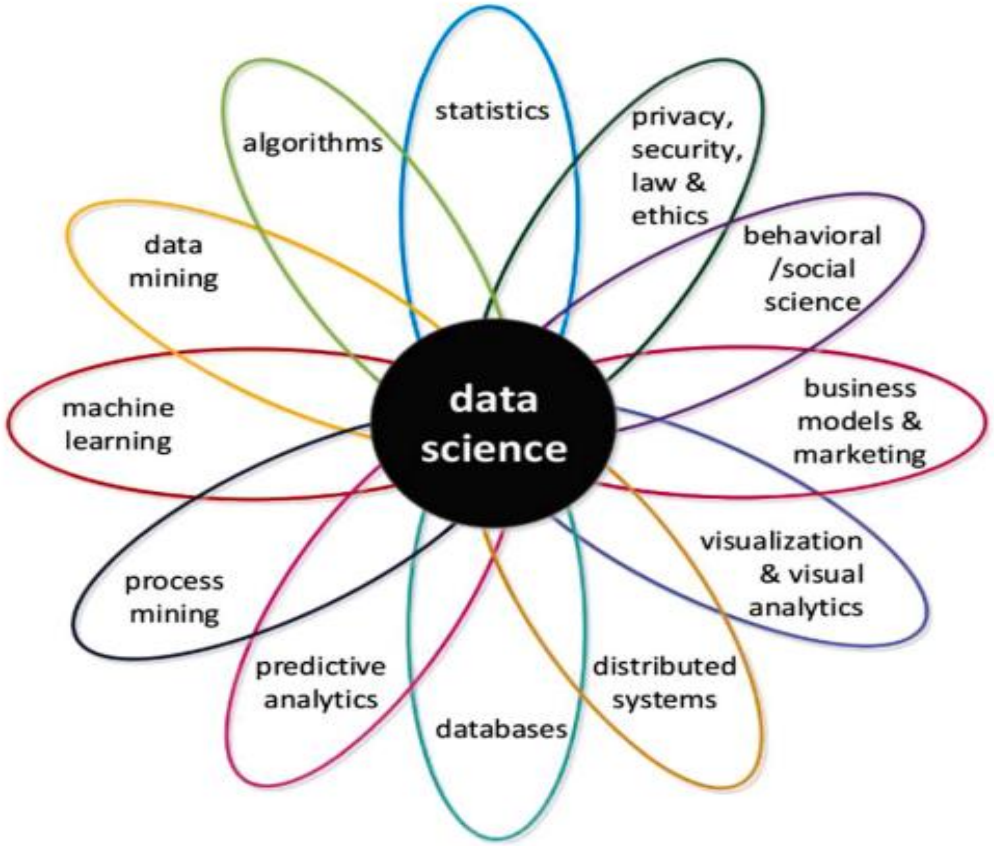


Fig. 2. Data science ingredients (van der Aalst 2016, p. 12).

- **Statistical Models vs. Machine Learning**:
  Statisticians tend to develop models confirmed by data.
  data scientists often focus on machine learning and data mining without being restricted by models

- **Job Market Trends**

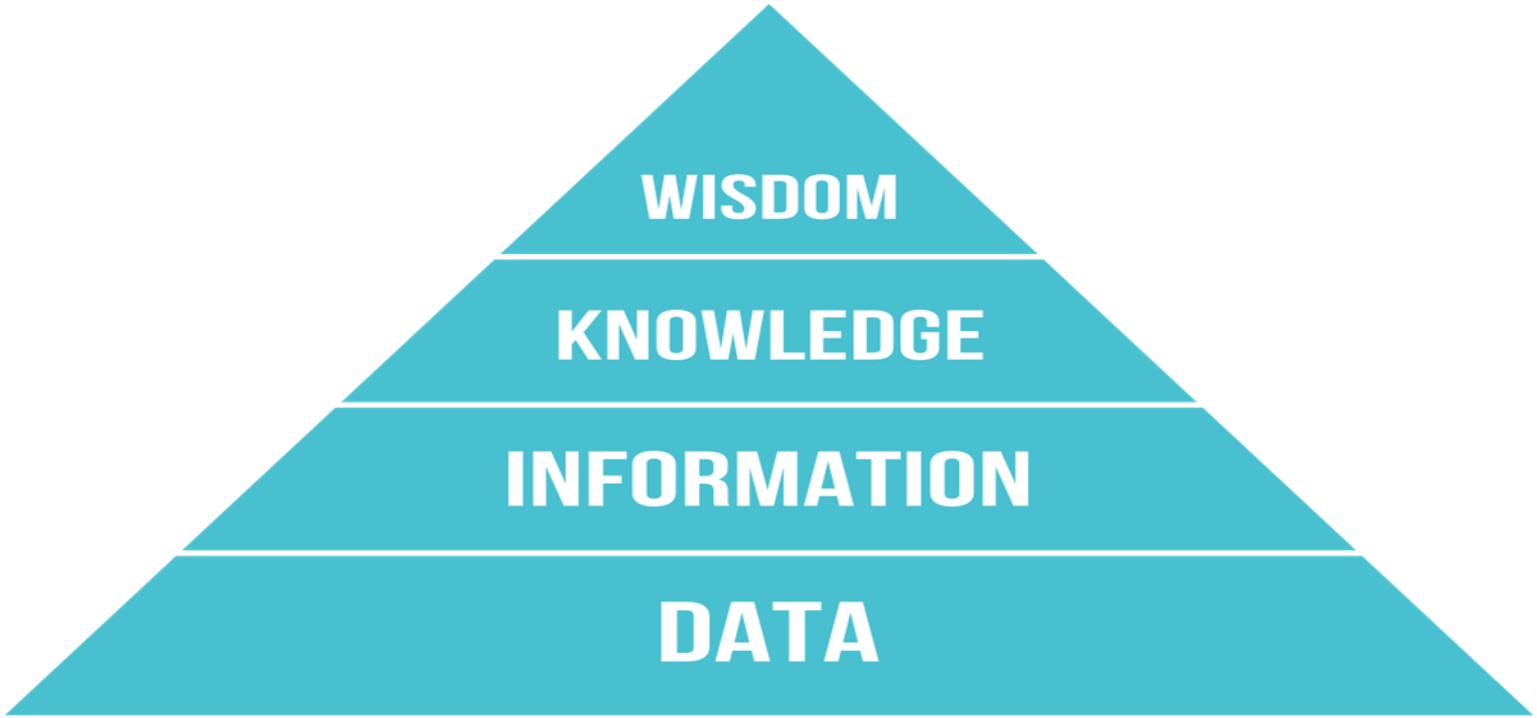
  google trends ***"whilst the role of a data scientist is increasingly more popular, as a field, the popularity of statistics continues to dominate over data science"***

- **Threat or Opportunity**
  Some view data science as a challenge, while others see it as a way for statistics to evolve and benefit.

- **Importance of Data**
  The recognition of the importance of data has led to increased demand for data scientists and statisticians who can interpret and work with data effectively.

The challenges posed by the evolution of big data

- **The Impact of Big Data**
  The development of data collection procedures, particularly through ambient intelligence, has significantly affected the characteristics of Big Data
- **The DIKW Pyramid**
  is used to understand the relationships within data



- **Combining Statistics and Data Science**

  Data scientists who combine statistical knowledge, machine learning, and analytics can provide valuable insights.
  Statistics can help ensure the mathematical correctness of data analytics.

- **SWOT Analysis**

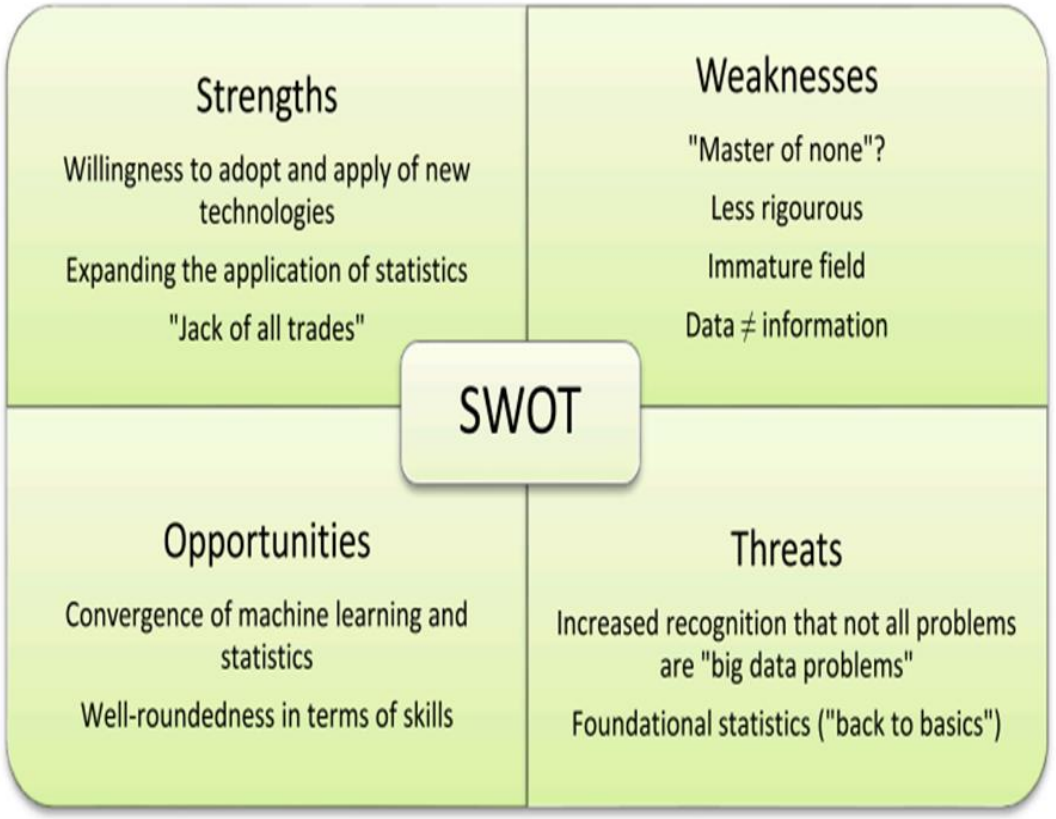


Fig. 4. SWOT analysis from the perspective of data science.

The limitations of data science

- **Challenges Faced by Data Scientists**:

- **Overcoming Barriers**

  Machine learning experts have explored strategies to overcome these barriers.
  One approach is k-fold cross-validation, which is popular but has limitations
  Another promising approach is "AutoML," where a machine learning algorithm is used to optimize the parameters of another algorithm.

- **Data Quality Concerns**

- **Accuracy vs. Interpretability**

  Data analytics involves a trade-off between accuracy and interpretability.

  Highly accurate machine learning models may lack interpretability.
  statistical methods are more interpretable but may sacrifice some accuracy.

- **Combining Statistics and Data Science**

  Data scientists who are knowledgeable in statistics and statisticians with data science skills are better positioned to navigate the trade-off between accuracy and interpretability

- **Data Signal and Noise**

  Big Data often contains a significant amount of noise, which can lead to misinterpretations.

- **SWOT Analysis from a Statistician's Perspective**

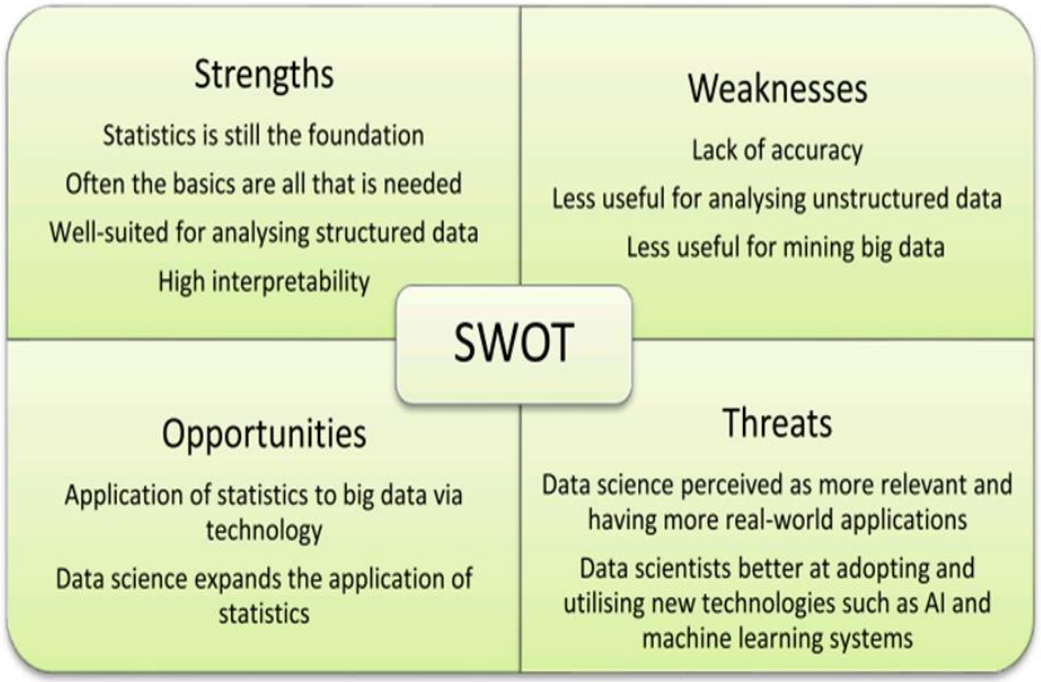


Fig. 6. SWOT analysis from the perspective of statistics.

Discussion

	DATA SCIENCE	STATISTICS
THEORITICAL ORIGINS	NEW FIELD	DEEP HISTORICAL ROOTS
FOCUS	PRACTICAL SOLUTION TO REAL WORLD PROBLEM	THEORITICAL SOPHISTICATION
APPROACHES	MACHINE LEARNING	EMPHASIS METHODOLOGY AND MODEL DEVELOPMENT
MODEL BUILDING	PREDECTION ERRORS AND FEATURES IDENTIFICATIONS	EXAMINING CORRELATIONS, CAUSALITY BETWEEN VARIABLES ,THEORY AND PREDICTORS
INTERPRETABILITY VS ACCURACY	PLACES A STRONGER EMPHASIS ON ACCURACY	EXCELS IN INTERPRETABILITY
DATA TYPE PREFERENCE	HIGHLY UNSTRUCTED DATA	WELL STRUCTED DATA


CONCLUSION