Introduction

Daniel Alexis Gutierrez Pachas, PhD
dgutierrezp@ucsp.edu.pe

Universidad Católica San Pablo | Departamento de Ciencia de la Computación

About me

Daniel Alexis Gutierrez Pachas, PhD
email: dgutierrezp@ucsp.edu.pe

Ph.D. in Computer Science and Computational Mathematics. I hold an undergraduate degree in Mathematics and a masters in Mathematics.

Actually, I am a full-time professor in Computer Science Department at the Universidad Católica San Pablo, Arequipa-Perú. Also, I have research experience in Optimization, Numerical Analysis, Stochastic systems and Data Science.

How much data is created every day?

41,666,667

messages shared
by WhatsApp users

1,388,889

video / voice calls made
by people worldwide

404,444

hours of video streamed
by Netflix users

347,222

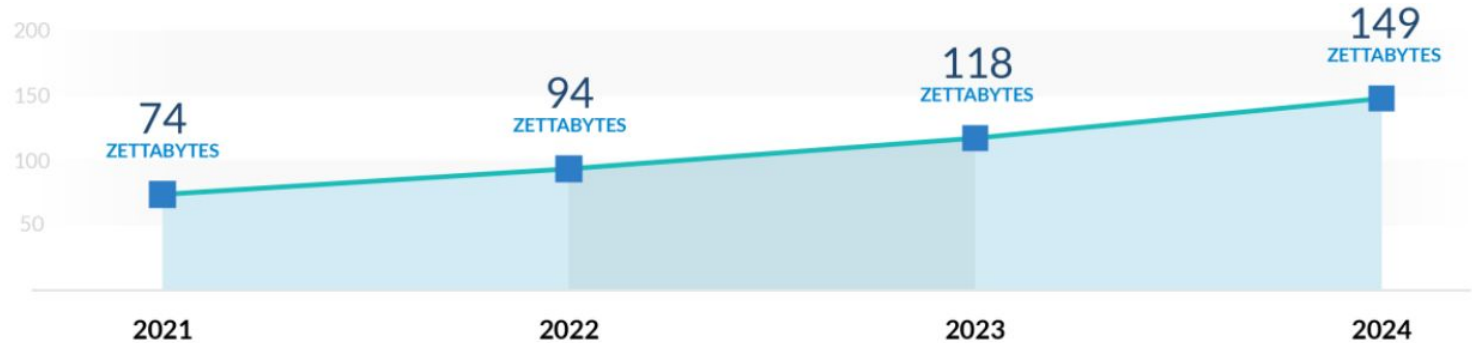stories posted by Instagram users

150,000

messages shared by Facebook users

147,000

photos shared by Facebook users

https://financesonline.com/how-much-data-is-created-every-day/

Estimated Data Consumption from 2021 to 2024

ZB = 10^{12} GB

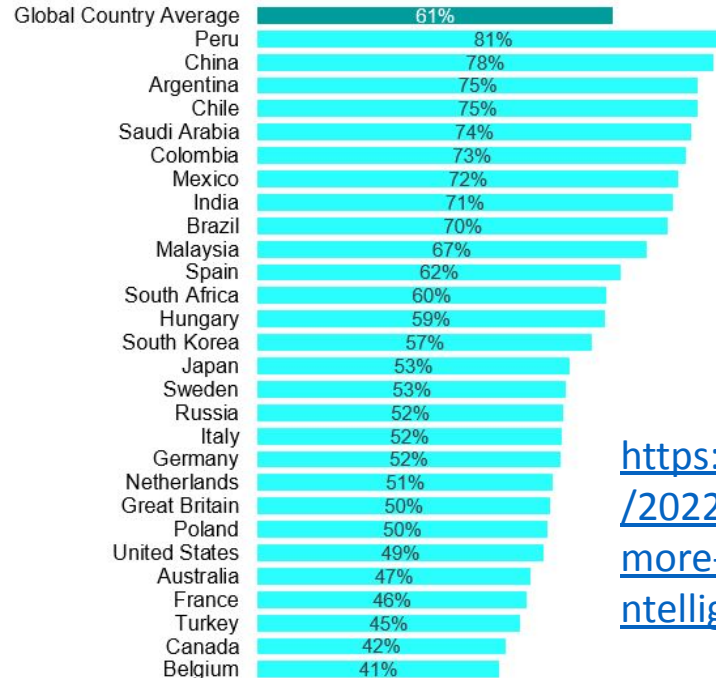https://financesonline.com/how-much-data-is-created-every-day/

OVERALL EXPECTATION OF LIFE **IMPROVEMENT** BECAUSE OF A.I. (AVERAGE FOR 13 AREAS)

Q. Please indicate whether you expect that the increased use of artificial intelligence will make each of the following better or worse for you and your family?

Average of % "Better" for 13 areas:
Education/learning new things
Safety
Employment
Shopping
Transportation
Entertainment
Cost of living
Home
Income
Environment
Food/nutrition
Personal and family relationships
Freedom/legal rights

Base: 19,504 online adults aged 16-74 across 28 countries, Nov.–Dec. 2021
Online samples in Brazil, Chile, mainland China, Colombia, India, Malaysia,
Mexico, Peru, Russia, Saudi Arabia, South Africa, and Turkey tend to be more
urban, educated, and/or affluent than the general population.
The "Global Country Average" reflects the average result for all the countries
and markets where the survey was conducted. It has not been adjusted to the
population size of each country or market and is not intended to suggest a total
result.

© Ipsos - Global Opinions and Expectations about AI - January 2022

Country	%
Global Country Average	61%
Peru	81%
China	78%
Argentina	75%
Chile	75%
Saudi Arabia	74%
Colombia	73%
Mexico	72%
India	71%
Brazil	70%
Malaysia	67%
Spain	62%
South Africa	60%
Hungary	59%
South Korea	57%
Japan	53%
Sweden	53%
Russia	52%
Italy	52%
Germany	52%
Netherlands	51%
Great Britain	50%
Poland	50%
United States	49%
Australia	47%
France	46%
Turkey	45%
Canada	42%
Belgium	41%

https://www.weforum.org/press/2022/01/emerging-economies-more-optimistic-about-artificial-intelligence-survey-finds

TRUST IN A.I. IS CORRELATED WITH PERCEIVED UNDERSTANDING; BOTH ARE HIGHER IN EMERGING COUNTRIES THAN IN HIGH-INCOME COUNTRIES
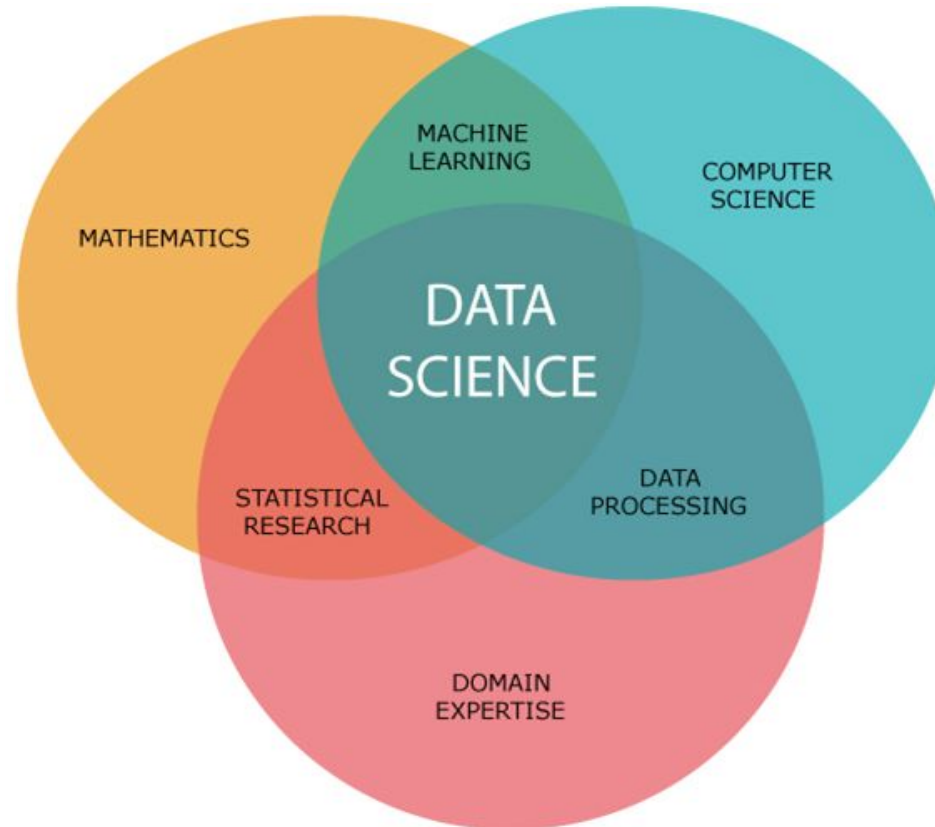
Components of Data Science

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21th century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Databases SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Universidad Católica **San Pablo** | **Departamento de Ciencia de la Computación**

Goals

- Introduce Statistical Fundamentals.

- Data Exploration using Python and R.

- Identify relevant findings and communicate them.