

Course Introduction

Il-Youp Kwak

High Level Goals?

Become a good **Data Scientist** with background in **Statistics**

Exponential Laws of Computing Growth

Moore's Law – The number of t

Moore's law describes the empirical regularity that the
This advancement is important as other aspects of tech
linked to Moore's law.

50,000,000,000

10,000,000,000

5,000,000,000

1,000,000,000

500,000,000

Transistor count

100.000.000

50 000 000

10.000.000

5.000.000

1,000,000

500,000

100,000

50,000

10,000

5.000

Intel 8088, IBM XT, 8bit

Intel Pentium, i486

286

386

486

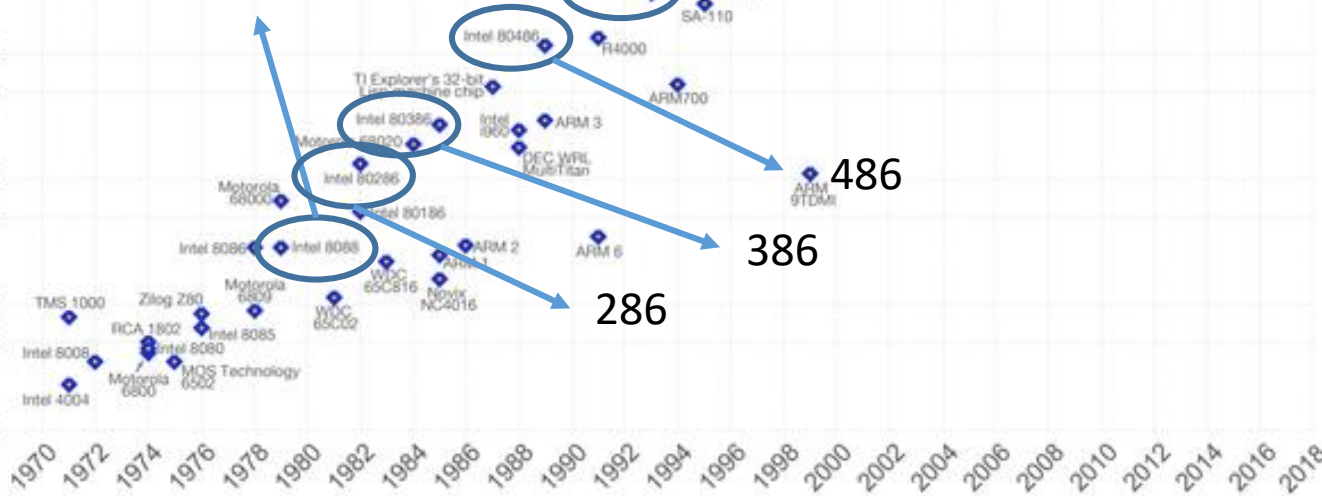
grated circuit chips (1971-2018)

rated circuits doubles approximately every two years.

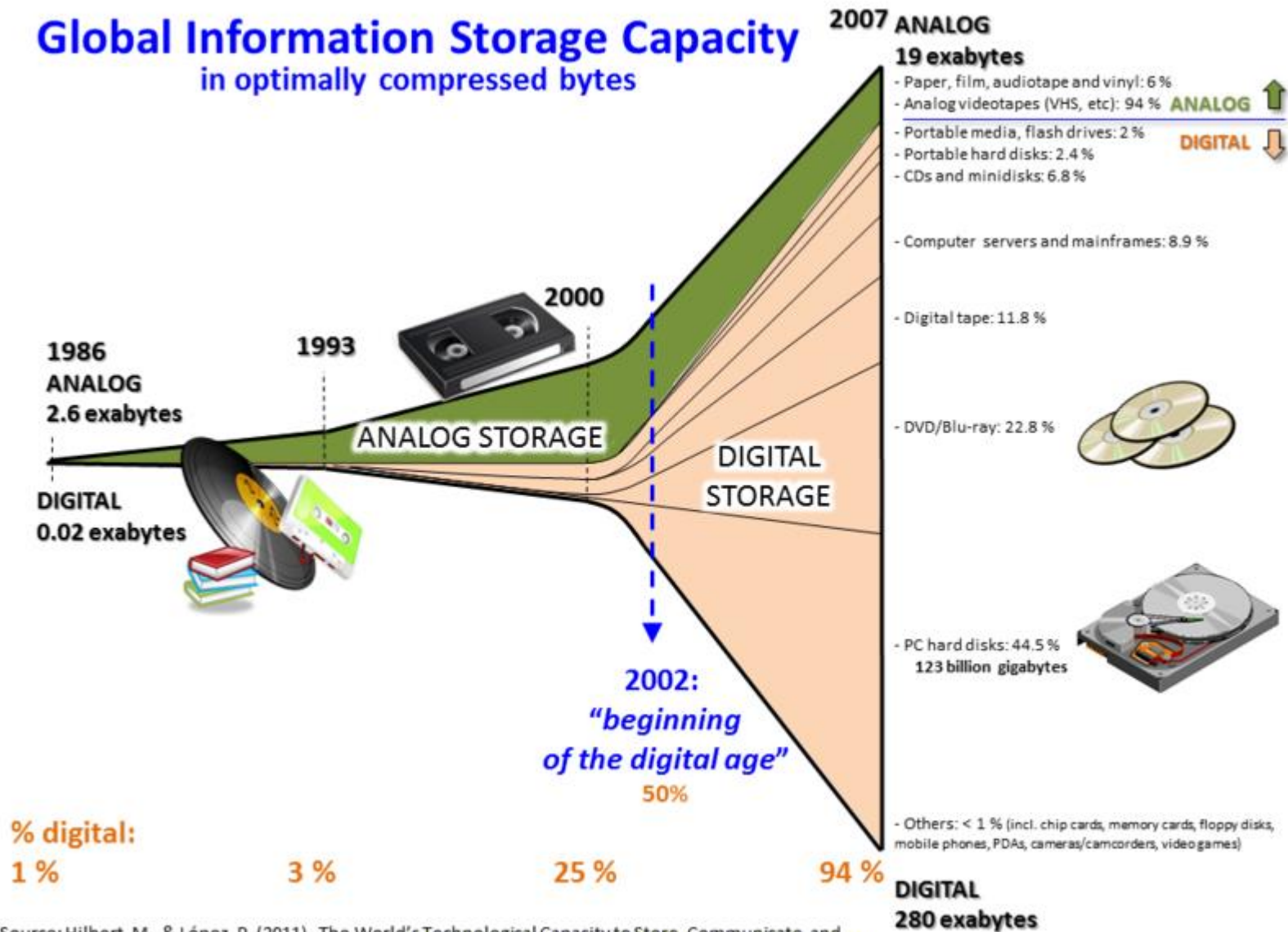
rocessing speed or the price of electronic products – are



Getting Multiple cores



Global Information Storage Capacity also Growing



Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60–65. <http://www.martinhilbert.net/WorldInfoCapacity.html>

Opens **Big Data** Era



A good eco for **AI**





BEST JOB IN AMERICA: DATA SCIENTIST!

Glassdoor recently ranked the 50 Best Jobs in America
3 of the top 5 are in the Data Analytics Field





#1 Data Scientist
4000+ openings on Glassdoor
\$110K median base salary
4.4/5.0 level of job satisfaction



#3 Data Engineer
2500+ openings on Glassdoor
\$106K median base salary
4.3/5.0 level of job satisfaction



#5 Analytics Manager
2000 openings
\$112K median base salary
4.1/5.0 level of job satisfaction

Glassdoor recently ranked the 50 Best Jobs in America
3 of the top 5 are in the Data Analytics Field

Glassdoor recently ranked the 50 Best Jobs in America
3 of the top 5 are in the Data Analytics Field



4000+ openings on Glassdoor
\$110K median base salary
4.4/5.0 level of job satisfaction

4000+ openings on Glassdoor
\$110K median base salary
4.4/5.0 level of job satisfaction

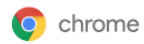
2500+ openings on Glassdoor
\$106K median base salary
4.3/5.0 level of job satisfaction

2500+ openings on Glassdoor
\$106K median base salary
4.3/5.0 level of job satisfaction

2000 openings
\$112K median base salary
4.1/5.0 level of job satisfaction

2000 openings
\$112K median base salary
4.1/5.0 level of job satisfaction

What they do with **Data**?



DOWNLOAD ▾

SET UP ▾

CHROMEBOOKS ▾

CHROMECAST ▾

Get a fast, free web browser

One browser for your computer, phone and tablet

[Download Chrome](#)

For Mac OS X 10.9 or later

[Download Chrome for another platform](#)



Google Chrome

Features

Do More with Chrome

Extensions

Enterprise



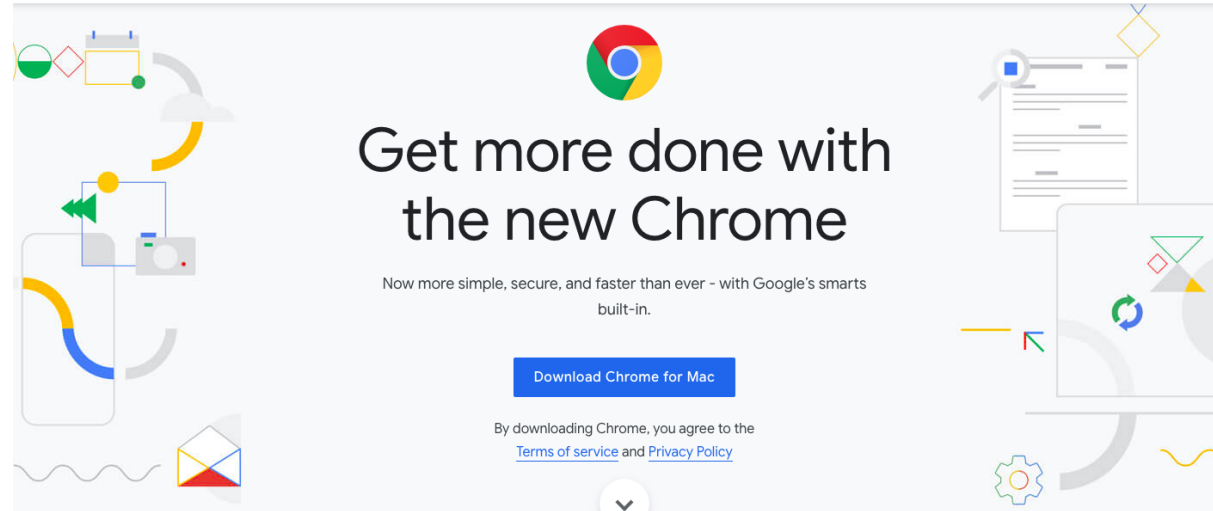
Get more done with the new Chrome

Now more simple, secure, and faster than ever - with Google's smarts built-in.

[Download Chrome for Mac](#)

By downloading Chrome, you agree to the

[Terms of service](#) and [Privacy Policy](#)



Searching System Objectives for e-commerce?

1. Profit



2. User Satisfaction



Let's discuss what we can do!

Skills we need

- 1. Computing skills**
- 2. Defining objectives and solving it**
- 3. Openness to learning new things**

How to become a Data Scientist?

기말 프로젝트 주제

1. 본인이 관심있는 빅데이터 선정
2. 풀고자 하는 문제 제시
3. 데이터 분석 및 발표

How you will look in presentation 😊



Email your goal to future you ☺

The image shows an email composition window. At the top, the recipient is '받는 사람' (받는 사람) with a profile picture and the name '곽일엽' (곽일엽). Below this is a '참조' (참조) button. The main body of the email contains the text 'My Goal'. At the bottom, there is a toolbar with various icons for editing and sending. A red box labeled '1' highlights the 'More' menu icon (three dots) in the toolbar. A second red box labeled '2' highlights the '예약발송' (예약발송) option in the dropdown menu. A third red box labeled '3' highlights the '예약 시간 지정' (예약 시간 지정) option in the sub-menu that appears after clicking '예약발송'.

받는 사람 곽일엽 ✕

참조

My Goal

초안 저장

서명 삽입

보낸 사람 표시

중요도 설정 >

일반 텍스트로 전환

접근성 문제 확인

암호화

메시지 옵션 표시...

중양대학교 조직도

표시이름변경

예약발송 >

예약발송

예약 시간 지정

예약 발송 목록

대용량파일첨부

추가 기능 가져오기

보내기 취소

...

Course Overview

1	과일업	Introduction, R vs Python	
2	과일업	Markdown, Github	
3	과일업	R for Data Science - dplyr, ggplot	HW 1
4	과일업	Python for big data and data handling	
5	과일업	Batch learning using python	HW 2
6	과일업	Neural Networks and Deeplearning	
7	과일업	Improving deep neural networks: hype rparameter tuning, regularization and optimization	
8	과일업	중간고사	

Course Overview

9	과일업	Convolutional Neural Networks 1	HW3
10	과일업	Convolutional Neural Networks 2	
11	과일업	Data Generator for handling big data	HW4
12	과일업	Transfer Learning	
13	과일업	deep learning for spoof detection 1	
14	과일업	deep learning for spoof detection 2	
15	과일업	기말 프로젝트 발표 및 토의 1	
16	과일업	기말 프로젝트 발표 및 토의 2	

R vs Python

R의 장점

- 좋은 통계 라이브러리가 많음
- 데이터 분석/시각화 용이

R의 단점

- 대용량 데이터 분석에 느릴 수 있음

Python의 장점

- Programing 언어로 자유도 높음
- 대용량 데이터 작업, 전처리에 좋음

Python의 단점

- 프로그래밍 언어에 대한 이해도 필요

Python 공부 (Jump to python)

Minimum 2장, 3장, 4장
Optional 5장, 7장

기말 발표주제 정해지면 메일로 알려줄것

Thank you!
Q & A