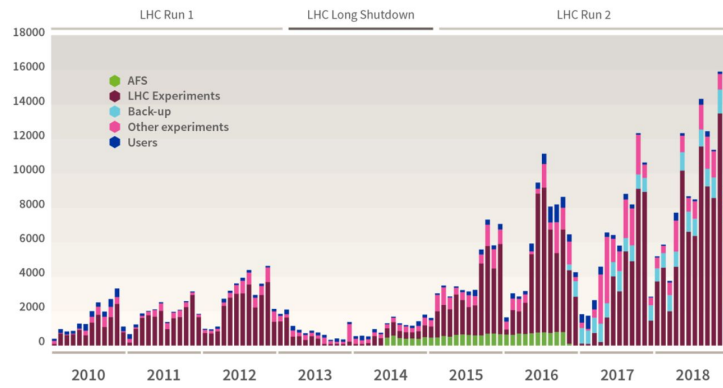
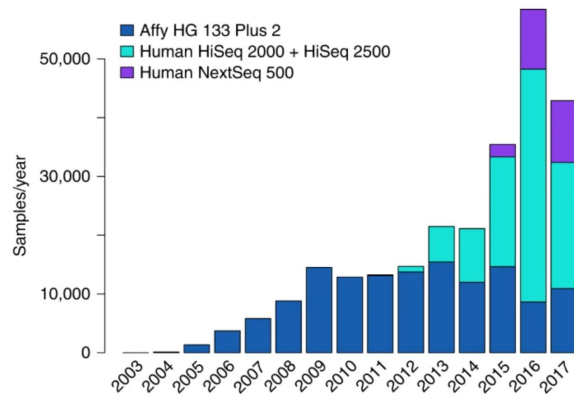


Software Engineering for Scientist



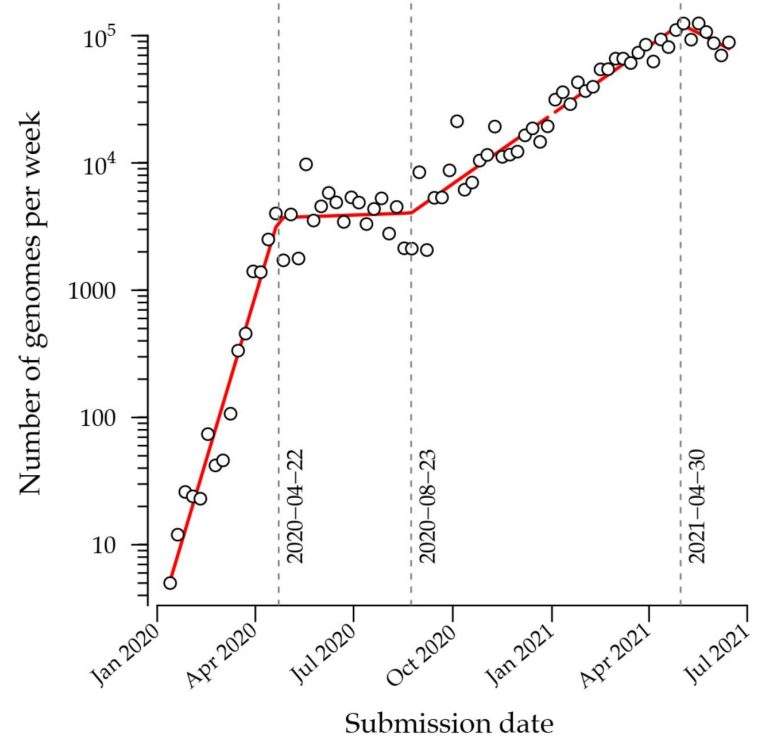
Data (in terabytes) recorded on tape at CERN month-by-month. This plot shows the amount of data recorded on tape generated by the LHC experiments, other experiments, various back-ups and users. In 2018, over 115 PB of data in total (including about 88 PB of LHC data) were recorded on tape, with a record peak of 15.8 PB in November (Image: Esma Mobs/CERN)

Ozcesmeci, 2019



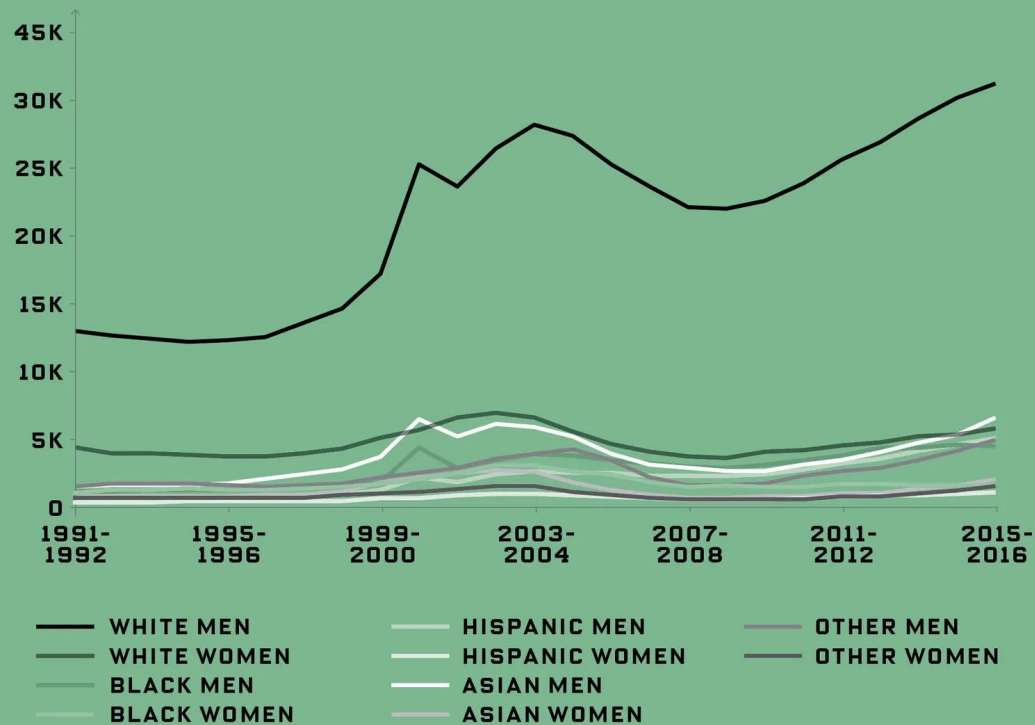
Publicly available RNA-seq samples currently available at GEO/SRA for human and mouse compared to available samples collected with the popular Affymetrix HG UI33 Plus 2 platform

Lachmann, 2018



Ferreira, 2021

The Sheer Volume of White Male Computer Science Majors Continues to Rise



SOURCE: CENTER FOR EDUCATION STATISTICS, INTEGRATED POSTSECONDARY EDUCATION DATA SYSTEM

Computer Science 1: Starting Computing

Computer Science 1: Starting Computing



Computer Science 2: Data Structures

Computer Systems

Discrete Structures

Software Development Methods and Tools

Principles of Programming Languages

Algorithms

Linear Algebra

Introduction to Data Science

CSCI Senior Capstone

Foundations of Software Engineering

Computer Science 1: Starting Computing



Computer Science 2: Data Structures

Computer Systems

Discrete Structures

Software Development Methods and Tools

Principles of Programming Languages

Algorithms

Linear Algebra

Introduction to Data Science

CSCI Senior Capstone

Foundations of Software Engineering

Computer Science 1: Starting Computing



Computer Science 2: Data Structures

Computer Systems

Discrete Structures

Software Development Methods and Tools

Principles of Programming Languages

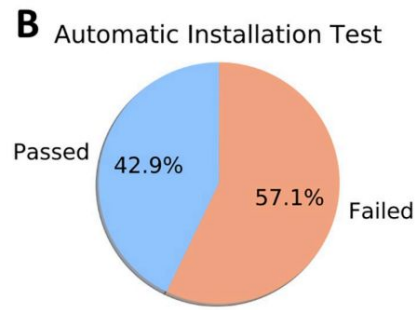
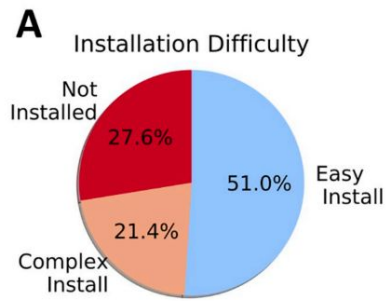
Algorithms

Linear Algebra

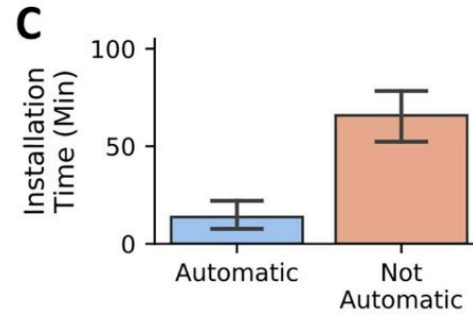
Introduction to Data Science

CSCI Senior Capstone

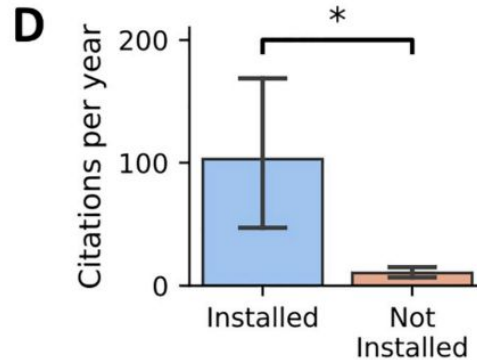
Foundations of Software Engineering



Installation is hard



and slow



and it is affects impact

A random forest-based framework for genotyping and accuracy assessment of copy number variations

Xuehan Zhuang, Rui Ye, Man-Ting So, Wai-Yee Lam, Anwarul Karim, Michelle Yu, Ngoc Diem Ngo, Stacey S Cherny, Paul Kwong-Hang Tam, Maria-Mercè Garcia-Barcelo, Clara Sze-man Tang, Pak Chung Sham

<https://github.com/sunnyzxh/CNV-JACG>

```
&showLog("Calculating het probability for $sample...");  
`/home/yerui/miniconda2/bin/perl $home/bin/Het-prob.pl -f $ref -m $snps -r $precnv -i $bam -o $outdir/$precnvbase.het.prob`;
```

Coding Best Practices
Version control
Git and GitHub
Testing
Continuous Integration
Code Reviews
Benchmarking
Workflows
Algorithms
Data Structures
Project



SWE4S