POLITECNICO
MILANO 1863

Gender Discrimination in Data Analysis:
a Socio-Technical Approach

Supervisor: Prof. Viola Schiaffonati   M.Sc. Thesis by:
Co-supervisor: Prof. Letizia Tanca   Riccardo Corona
                      Prof. Pierre Senellart   927975
                      Prof. Karine Gentelet   07/10/2021

Data analysis

Set of processes for inspecting, cleaning, transforming, and modeling data with the aim of discovering useful information, informing conclusions, and supporting decision making.

Gender discrimination

Specific (sub)category of social problems, here expressed in the form of the so-called '**gender gap**', definable as:

A difference between the way men and women are treated in society, or between what men and women do and achieve.

Problem

Data and datasets, on which a lot of actions of our daily routine are based, can be **unfair**. Unfair, or better to say, **biased** data, may influence, directly or indirectly, our perception of reality, and lead us to make decisions that, although seemingly fair and just, contain in turn bias, and discriminate against individuals or groups of individuals.
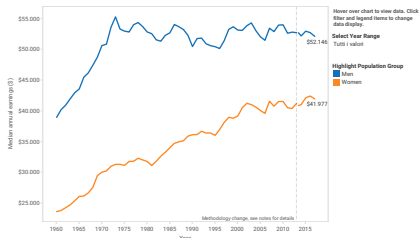
Example scenarios

- COMPAS tool used in the U.S. to predict recidivism risk biased against Black people (2016).
- Amazon software to screen candidates for employment biased against women (2015).

Median annual earnings by sex
March 1960-2017

Hover over chart to view data. Click filter and legend items to change data display.

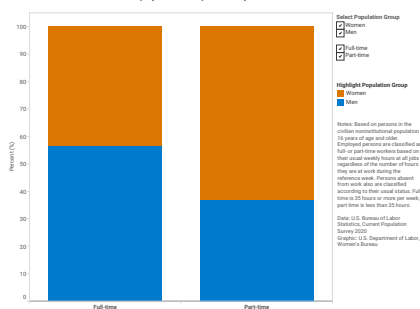**Select Year Range**
Tutti i valori

Highlight Population Group
- Men
- Women

Notes: Earnings are based on median annual earnings of full-time, year-round workers, 15 years old and over as of March of the following year for previous years. Before 1989 earnings are for civilian workers only.
The comparability of historical data has been affected at various times by methodological and other changes in the Current Population Survey. The 2014 CPS ASEC included redesigned questions for income and health insurance coverage for a subsample of the 98,000 addresses using a probability split panel design. Approximately 68,000 addresses were eligible to receive a set of income questions similar to those used in the 2013 CPS ASEC and the remaining 30,000 addresses were eligible to receive the redesigned income questions, resulting in two estimates for 2013. Estimates based on the portion of the sample that received the redesigned income questions are the most appropriate for comparing estimates from ASEC 2014 with ASEC 2015 and beyond.
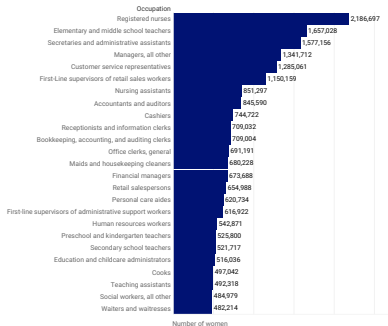Earnings are in 2017 CPI-U-RS adjusted dollars.
Source: 1961-2018 Annual Social and Economic Supplements, Current Population Survey, U.S. Census Bureau.
Graph by the Women's Bureau, U.S. Department of Labor

Percent distribution of workers employed full- and part-time by sex

Select Population Group
- ☑ Women
- ☑ Men
- ☑ Full-time
- ☑ Part-time

Highlight Population Group
- Women
- Men

Notes: Based on persons in the civilian noninstitutional population 16 years of age and older. Employed persons are classified as full- or part-time workers based on their usual hours at all jobs regardless of the number of hours they are at work during the reference week. Persons absent from work also are classified according to their usual status. Full time is 35 hours or more per week; part time is less than 35 hours.

Data: U.S. Bureau of Labor Statistics, Current Population Survey 2020
Graphic: U.S. Department of Labor, Women's Bureau

Most Common Occupations for Women in the Labor Force

Occupation	Number of women
Registered nurses	2,186,697
Elementary and middle school teachers	1,657,028
Secretaries and administrative assistants	1,577,156
Managers, all other	1,341,712
Customer service representatives	1,285,061
First-Line supervisors of retail sales workers	1,150,159
Nursing assistants	851,297
Accountants and auditors	845,590
Cashiers	744,722
Receptionists and information clerks	709,032
Bookkeeping, accounting, and auditing clerks	709,004
Office clerks, general	691,191
Maids and housekeeping cleaners	680,228
Financial managers	673,688
Retail salespersons	654,988
Personal care aides	620,734
First-line supervisors of administrative support workers	616,922
Human resources workers	542,871
Preschool and kindergarten teachers	525,800
Secondary school teachers	521,717
Education and childcare administrators	516,036
Cooks	497,042
Teaching assistants	492,318
Social workers, all other	484,979
Waiters and waitresses	482,214

Note: Full-time, year-round civilian employed 16 years and older. Occupations with at least 100 sample observations.
Data: U.S. Census Bureau, American Community Survey 2019
Graphic: U.S. Department of Labor, Women's Bureau.

Percentage of science, technology, engineering, and math (STEM) workers who are women
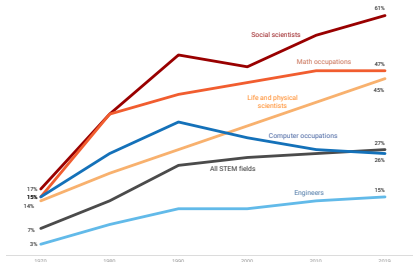
Note: STEM occupations are classified according to the Standard Occupational Classification STEM recommendations for presentation of government data available at
https://www.bls.gov/soc/Attachment_C_STEM_2018.pdf
Source: U.S. Census Bureau, decennial census 1970-2000 and American Community Survey public use microdata 2010 and 2019.
Graphic by the Women's Bureau, U.S. Department of Labor

Section 1

Riccardo Corona

POLITECNICO MILANO 1863

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Subsection 1.1

This frame has an empty title.

- item 1
  - item 1.1
  - item 1.2
- item 2
- item 3

Riccardo Corona

POLITECNICO MILANO 1863

Slide 1.2 without numbering

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Block

Text.

Block
Text.

Alert block
Alert text.

Riccardo Corona

POLITECNICO MILANO 1863

Block

Text.

Alert block

Alert text.

Example block

Example text.

Riccardo Corona

POLITECNICO MILANO 1863