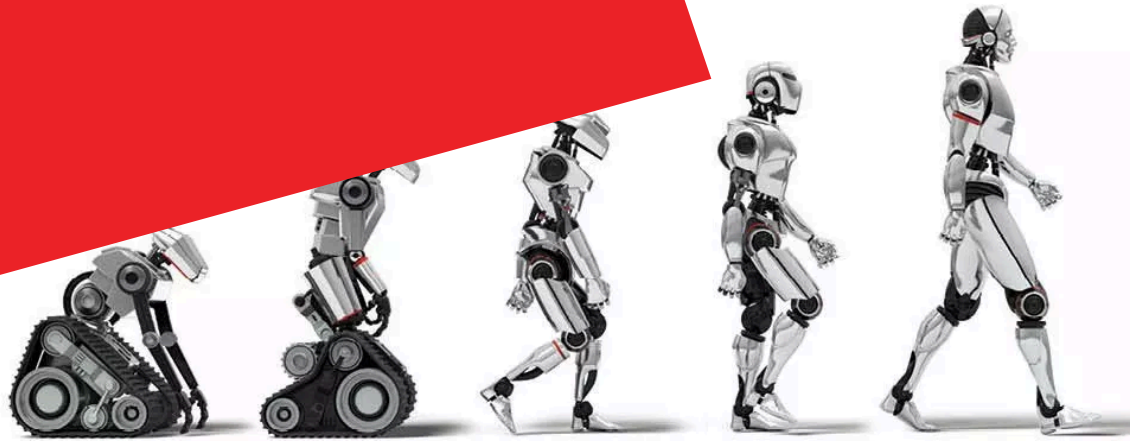
Using Big Data to Understand the Workforce
How Will ML Transform the Economy?

Erik Brynjolfsson
MIT (coming to Stanford July 2020!)

NSF Workshop
May 19, 2020

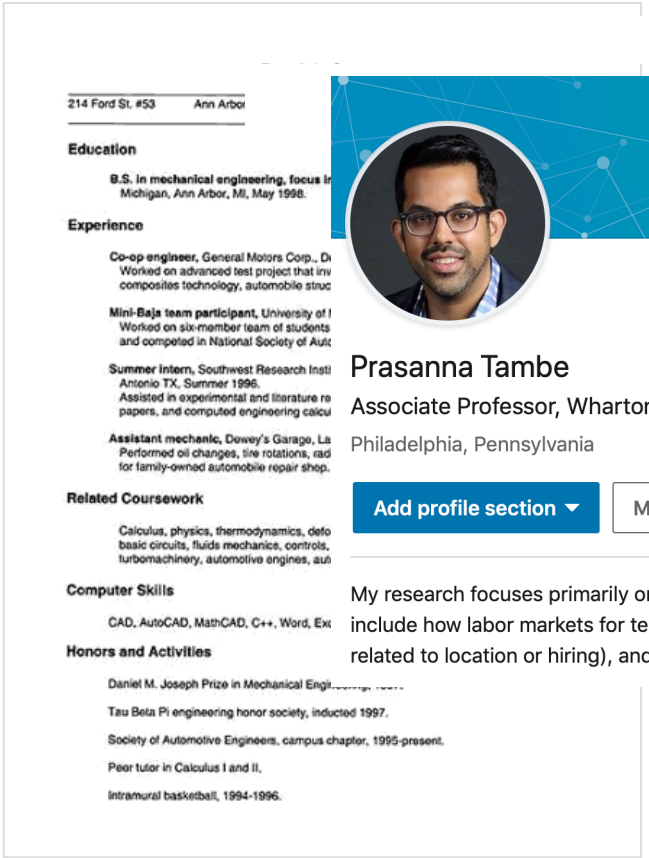


New Tools Beget Revolutions

Big Data is a Measurement Revolution

- Job Postings
- LinkedIn Profiles
- Mobile phone/GPS/Location data
- Google/Bing Searches
- Twitter feeds
- Clickstream/Page views/Web transactions
- Web links/Blog references/Facebook
- Email messages
- ERP/CRM/SCM transactions
- RFID (Radio Frequency Identification), Bar Code Scanner Data
- Real-time machinery diagnostics/engines/equipment
- Stock market transactions
- Wikipedia updates
- Etc….

Employment histories posted online provide rich information about firms and workers



214 Ford St. #53 Ann Arbo

Education

B.S. in mechanical engineering, focus I
Michigan, Ann Arbor, MI, May 1998.

Experience

Co-op engineer, General Motors Corp., D
Worked on advanced test project that inv
composites technology, automobile struc

Mini-Baja team participant, University of
Worked on six-member team of students
and competed in National Society of Auto

Summer intern, Southwest Research Insti
Antonio TX, Summer 1996.
Assisted in experimental and literature re
papers, and computed engineering calcu

Assistant mechanic, Dewey's Garage, La
Performed oil changes, tire rotations, rad
for family-owned automobile repair shop.

Related Coursework

Calculus, physics, thermodynamics, defo
basic circuits, fluids mechanics, controls,
turbomachinery, automotive engines, aut

Computer Skills

CAD, AutoCAD, MathCAD, C++, Word, Ex

Honors and Activities

Daniel M. Joseph Prize in Mechanical Engi

Tau Beta Pi engineering honor society, inducted 1997.

Society of Automotive Engineers, campus chapter, 1995-present.

Peer tutor in Calculus I and II.

Intramural basketball, 1994-1996.

Prasanna Tambe

Associate Professor, Wharton School @ U.Penn (OIDD)

Philadelphia, Pennsylvania

Add profile section

More...

The Wharton School

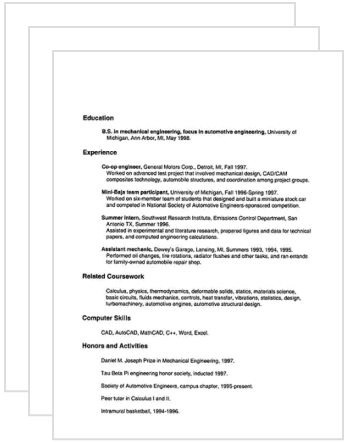
University of Pennsylvania - The Wharton School

See contact info

See connections (500+)

My research focuses primarily on the economics of technology and labor. Specific research interests include how labor markets for technical skills (e.g. data science) impact corporate decisions (e.g. related to location or hiring), and understanding the labor market for AI skills. Much of this research u...

Online employment databases: Employment histories for millions of US workers



Education

B.S. in mechanical engineering, focus in automotive engineering, University of Michigan, Ann Arbor, MI, May 1998.

Experience

Co-op engineer, General Motors Corp., Detroit, MI, Fall 1997.
Worked on advanced test project that involved mechanical design, CAD/CAM composites technology, automobile structures, and coordination among project groups.

Mini-Baja team participant, University of Michigan, Fall 1996-Spring 1997.
Worked on six-member team of students that designed and built a miniature stock car and competed in National Society of Automotive Engineers-sponsored competition.

Summer intern, Southwest Research Institute, Emissions Control Department, San Antonio TX, Summer 1996.
Assisted in experimental and literature research, prepared figures and data for technical papers, and computed engineering calculations.

Assistant mechanic, Davey's Garage, Lansing, MI, Summers 1993, 1994, 1995.
Performed oil changes, tire rotations, radiator flushes and other tasks, and ran errands for family-owned automobile repair shop.

Related Coursework

Calculus, physics, thermodynamics, deformable solids, statics, materials science, basic circuits, fluids mechanics, controls, heat transfer, vibrations, statistics, design, turbomachinery, automotive engines, automotive structural design.

Computer Skills

CAD, AutoCAD, MathCAD, C++, Word, Excel.

Honors and Activities

Daniel M. Joseph Prize in Mechanical Engineering, 1997.

Tau Beta Pi engineering honor society, inducted 1997.

Society of Automotive Engineers, campus chapter, 1995-present.

Peer tutor in Calculus I and II.

Intramural basketball, 1994-1996.

Leading online job search site provided resumes for about 40 million workers including fielded data

150 million employer-employee combinations

Can step backward through employment histories to create longitudinal measures

Employee Data		
Employee	Education	Occupation
Employee 1	4 Years College	IT
Employee 2	4 Years College	Sales

Employee Work History Data				
Employee	Employer Name	Job Title	Start Date	End Date
Employee 1	Firm Name 3	Project Manager	5-01-2006	Present
Employee 1	Firm Name 2	Software Engineer	9-01-2003	3-15-2006
Employee 2	Firm Name 2	Director of Technology	4-01-2006	Present
Employee 2	Firm Name 1	MIS Manager	1-01-2006	3-20-2006

nature
International weekly journal of science

Track how technology is transforming work

Without data on how artificial intelligence is affecting jobs, policymakers will fly blind into the next industrial revolution, warn Tom Mitchell and Erik Brynjolfsson.

Science

POLICY FORUM

TECHNOLOGY AND THE ECONOMY

What can machine learning do? Workforce implications

Profound change is coming, but roles for humans remain

By Erik Brynjolfsson[1,2] and Tom Mitchell[3]

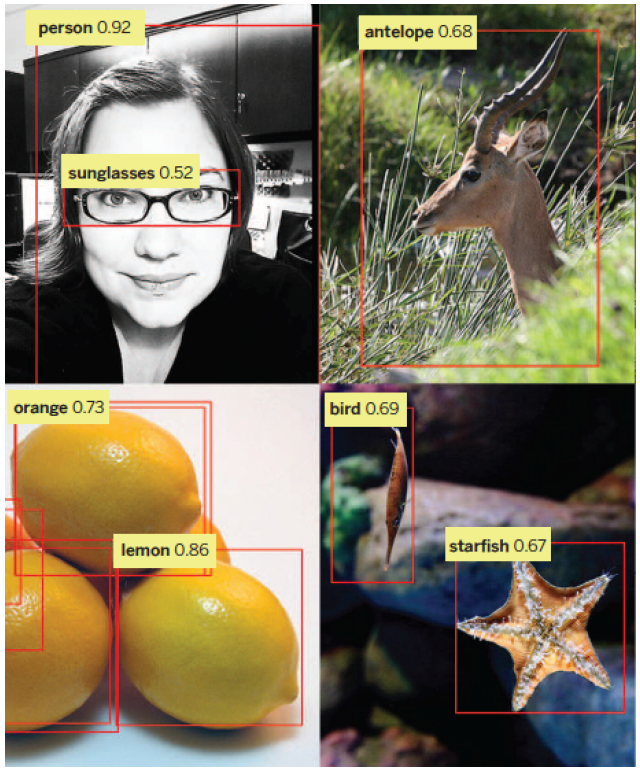
Digital computers have transformed work in almost every sector of the economy over the past several decades (*1*). We are now at the beginning of an even larger and more rapid transformation due to recent advances in machine learning (ML), which is capable of accelerating the pace of automation itself. However, although it is clear that ML is a "general purpose technology," like the steam engine and electricity, which spawns a plethora of additional innovations and capabilities (*2*), there is no widely shared agreement on the tasks where ML systems excel, and thus little agreement on the specific expected impacts on the workforce and on the economy more broadly. We discuss what we see to be key implications for the workforce, drawing on our rubric of what the current generation of ML systems can and cannot do [see the supplementary materials (SM)]. Although parts of many jobs may be "suitable for ML" (SML), other tasks within these same jobs do not fit the criteria for ML well; hence, effects on employment are more complex than the simple replacement and substitution story emphasized by some. Although economic effects of ML are relatively limited today, and we are not facing the imminent "end of work" as is sometimes proclaimed, the implications for the economy and the workforce going forward are profound.

Any discussion of what ML can and cannot do, and how this might affect the economy, should first recognize two broad, underlying considerations. We remain very far from artificial general intelligence (*3*). Machines cannot do the full range of tasks that humans can do (*4*). In addition, although innovations

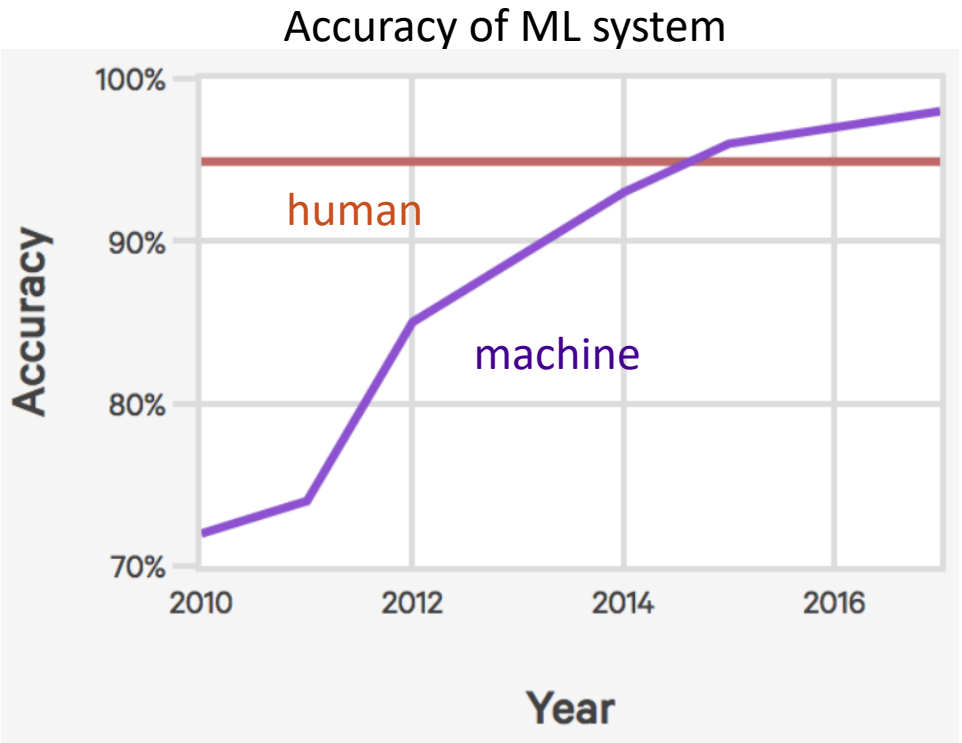
[1]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]National Bureau of Economic Research, Cambridge, MA 02138, USA. [3]Carnegie Mellon University, Pittsburgh, PA 15213, USA. Email: erikb@mit.edu

Published by AAAS

We've Crossed a Key Threshold



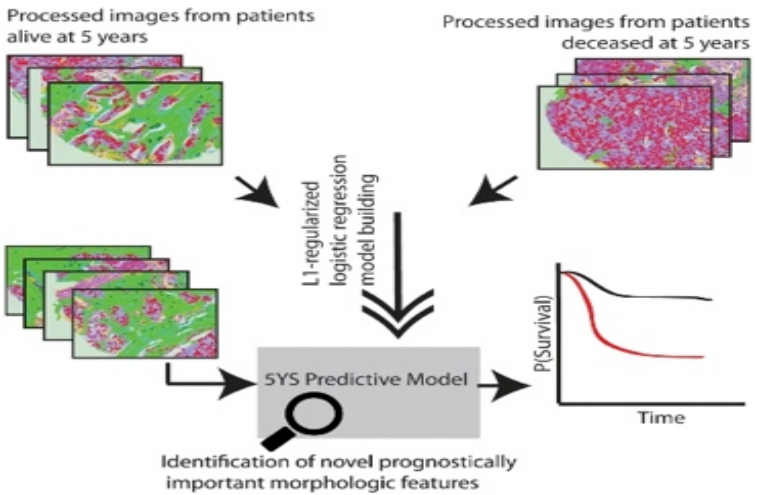
ImageNet Visual Recognition Challenge



source: The AI Index http://aiindex.org/

Problem Solving

Most of the recent progress in machine learning involves mapping from a set of inputs to a set of outputs



INPUT X	OUTPUT Y	APPLICATION
Voice recording	Transcript	Speech recognition
Historical market data	Future market data	Trading bots
Photograph	Caption	Image tagging
Drug chemical properties	Treatment efficacy	Pharma R&D
Store transaction details	Is the transaction fraudulent?	Fraud detection
Recipe ingredients	Customer reviews	Food recommendations
Purchase histories	Future purchase behavior	Customer retention
Car locations and speed	Traffic flow	Traffic lights
Faces	Names	Face recognition

Source: Brynjolfsson, Erik, and Andrew McAfee. "The Business of Artificial Intelligence." *Harvard Business Review* (2017).

Which Tasks Will Be Done by Machine Learning?

ML is far from AGI

We create a “Suitability for Machine Learning” (SML) rubric to assess tasks

- We apply it to 2,059 Detailed Work Activities in O*NET, 18,112 occupation-specific tasks, and 950 occupations (weighted by task importance)
- Questions are rated on five point scale from “strongly disagree” to “strongly agree”
- Each DWA is scored by 10 different people

Brynjolfsson, Mitchell and Rock, “What Can Machines Learn and What Does It Means for Occupations and the Economy, *AEA P&P*, 2018.

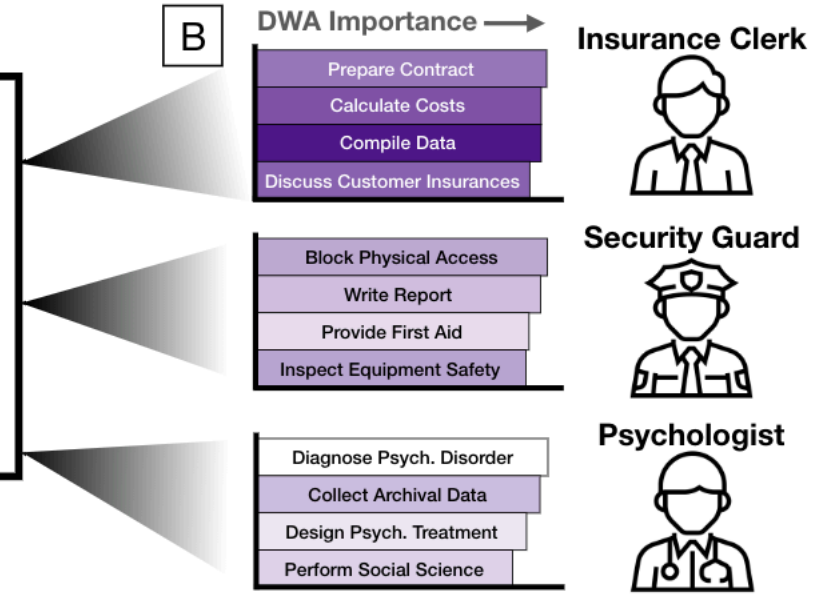
Use the Rubric to Evaluate ML Potential Impact

A

Detailed Work Activities (DWAs) are assessed for their suitability for machine learning:

- Does the task require complex reasoning?
- Is the task repeated frequently?
- Is the task predictive?
- Does the task get easier with practice?
- Is the task error tolerant?
- Does the task get easier with practice?
- Does the task rely on speech or visual data?

B



O*Net: Tasks Done by Radiologists (27 tasks)

Sample Tasks (out of 27 tasks):

1. Provide advice on types or quantities of radiology equipment needed to maintain facilities.
2. Perform interventional procedures such as image-guided biopsy, percutaneous transluminal angioplasty, transhepatic biliary drainage, or nephrostomy catheter placement.
3. Administer or maintain conscious sedation during and after procedures.
4. Interpret images using computer-aided detection or diagnosis systems.
5. Develop treatment plans for radiology patients.
6. Treat malignant internal or external growths by exposure to radiation from radiographs (x-rays), high energy sources, or natural or synthetic radioisotopes.
7. Conduct physical examinations to inform decisions about appropriate procedures.

O*Net: Tasks Done by Radiologists (27 tasks)

Sample Tasks (out of 27 tasks):

1. Provide advice on types or quantities of radiology equipment needed to maintain facilities.
2. Perform interventional procedures such as image-guided biopsy, percutaneous transluminal angioplasty, transhepatic biliary drainage, or nephrostomy catheter placement.
3. Administer or maintain conscious sedation during and after procedures.
4. **Interpret images using computer-aided detection or diagnosis systems.**
5. Develop treatment plans for radiology patients.
6. Treat malignant internal or external growths by exposure to radiation from radiographs (x-rays), high energy sources, or natural or synthetic radioisotopes.
7. Conduct physical examinations to inform decisions about appropriate procedures.

O*Net: Tasks Done by Radiologists (27 tasks)

Sample Tasks (out of 27 tasks):

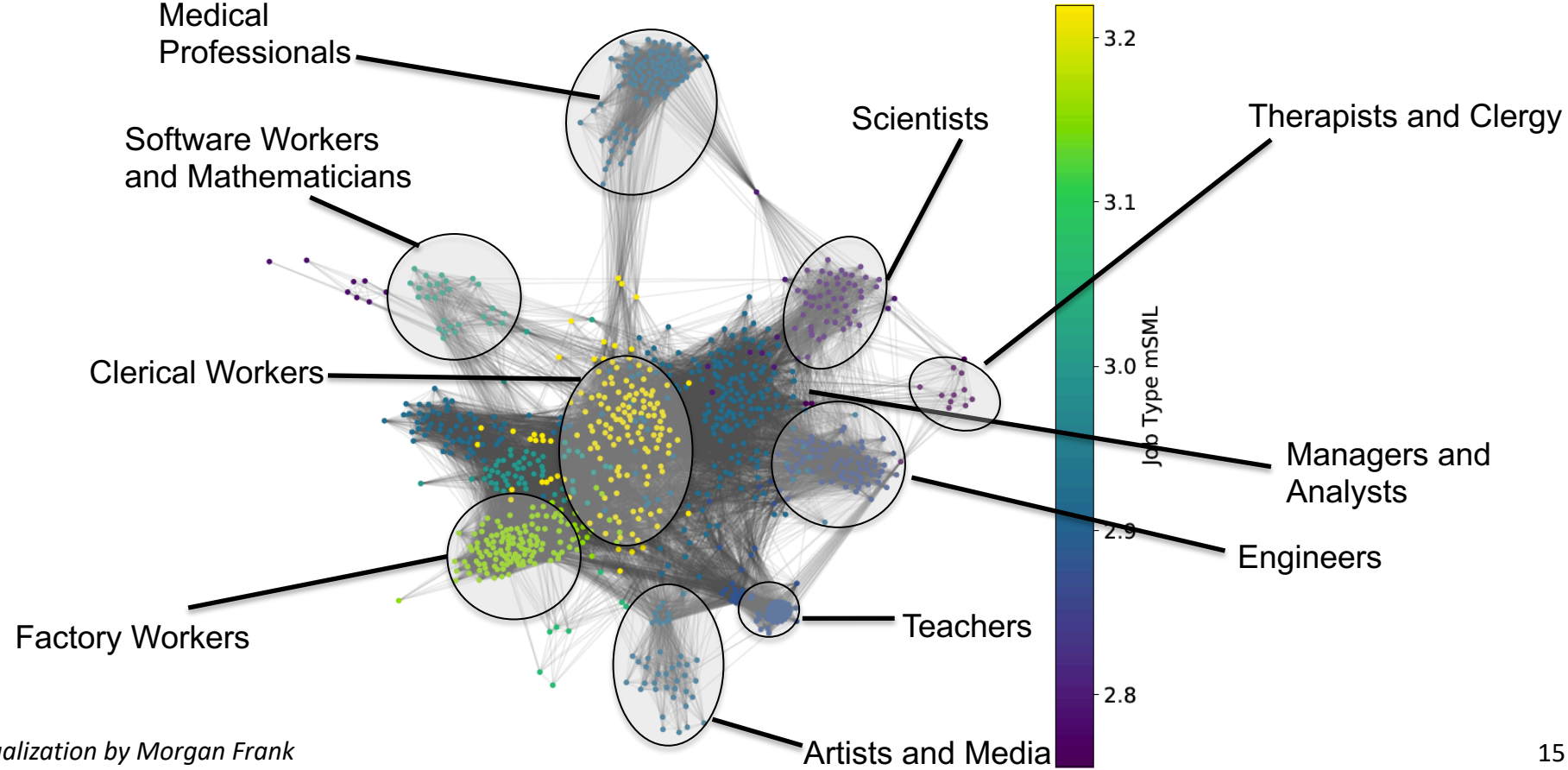
1. Provide advice on types or quantities of radiology equipment needed to maintain facilities.
2. Perform interventional procedures such as image-guided biopsy, percutaneous transluminal angioplasty, transhepatic biliary drainage, or nephrostomy catheter placement.
3. **Administer or maintain conscious sedation during and after procedures.**
4. **Interpret images using computer-aided detection or diagnosis systems.**
5. Develop treatment plans for radiology patients.
6. Treat malignant internal or external growths by exposure to radiation from radiographs (x-rays), high energy sources, or natural or synthetic radioisotopes.
7. Conduct physical examinations to inform decisions about appropriate procedures.

Total wage value of high SML, high Measurability Activities is $713 Billion

- Method: Take the task weights supplied by O*NET and multiply them by the occupational *wage*.
  - This is the *wage attributable to the task*
  - Calculate the *average wage attributable to the task over occupations*
  - Sum the wage bill attributable to the task over all occupations
    - This is the total wage bill in a given task (or activity)

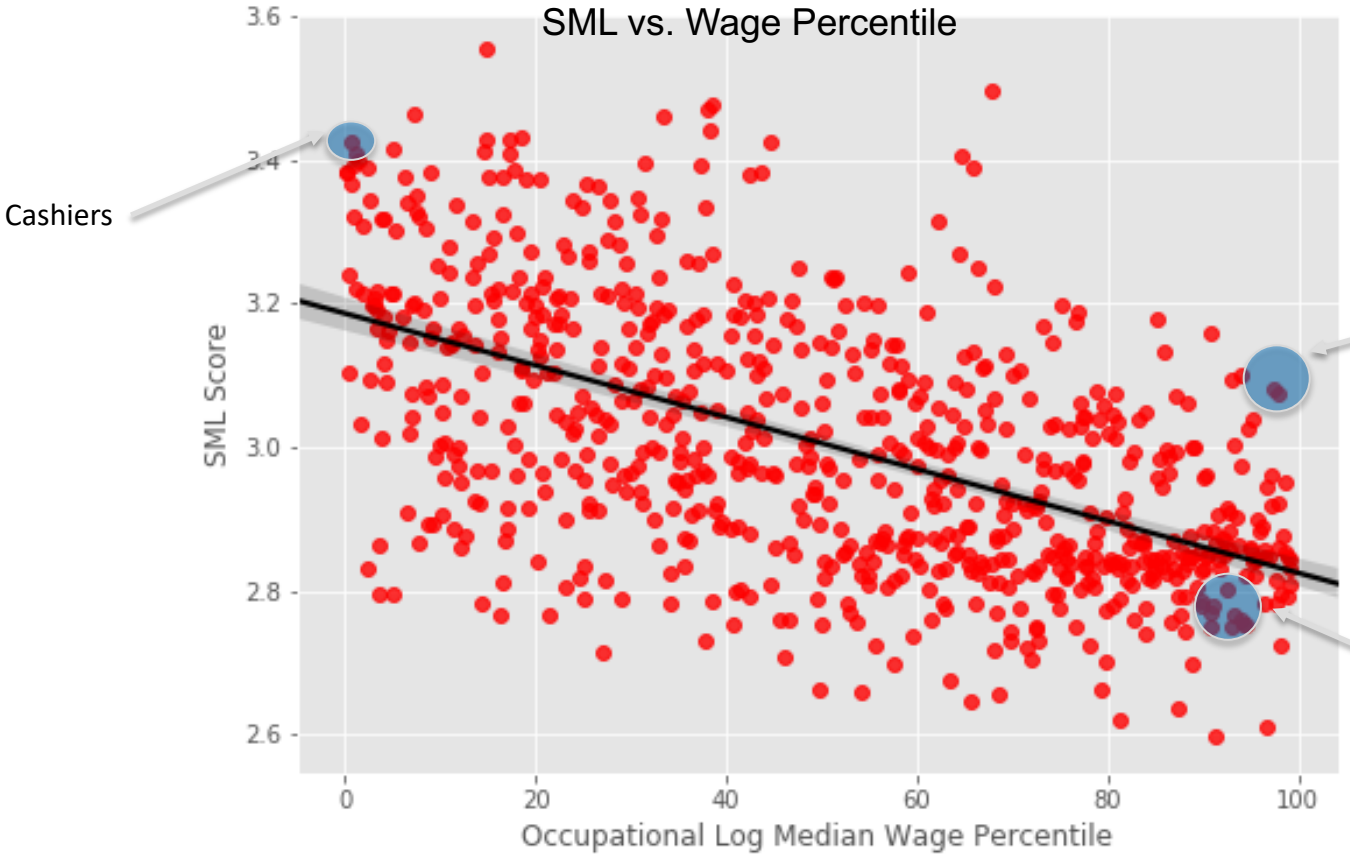
- Also calculate the high SML wage value, Job-specific wage value
  - *High SML, High Measurability wage value:* wage attributable to SML Tasks that are >90th Percentile and >4 Measurability
    - This total is $713 billion
  - *Job-specific low SML wage proportion*: (Value of low SML Activity in Job / Value of those Activities in overall economy)

Mapping to Jobs Connected by Activities



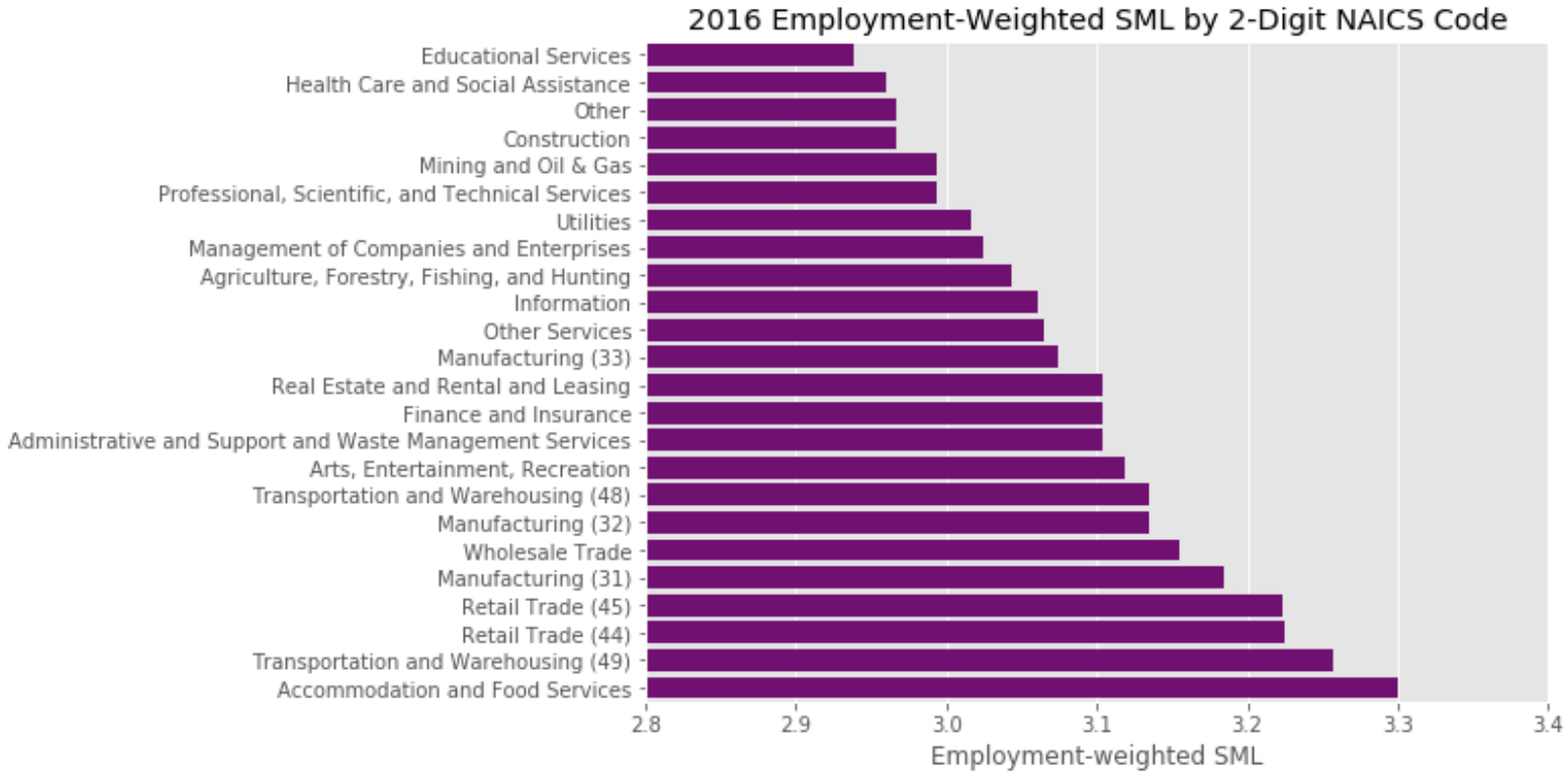
Visualization by Morgan Frank

ML will affect all groups, but especially lower wage workers

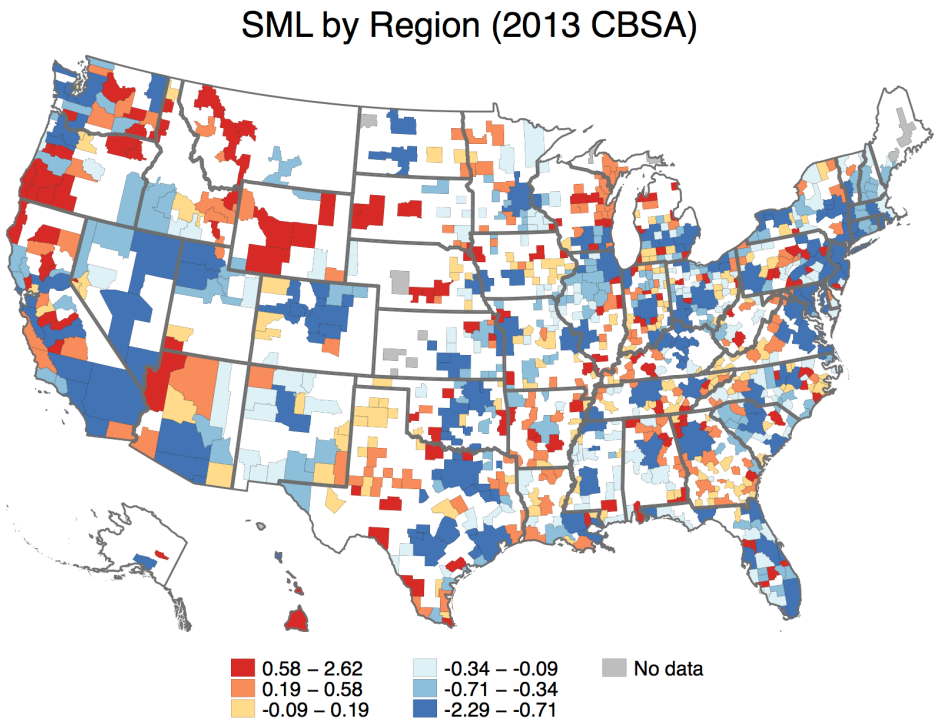


Regression of SML on Wages
Slope: -0.0034, t-stat: 18.5

How does SML vary by industry?



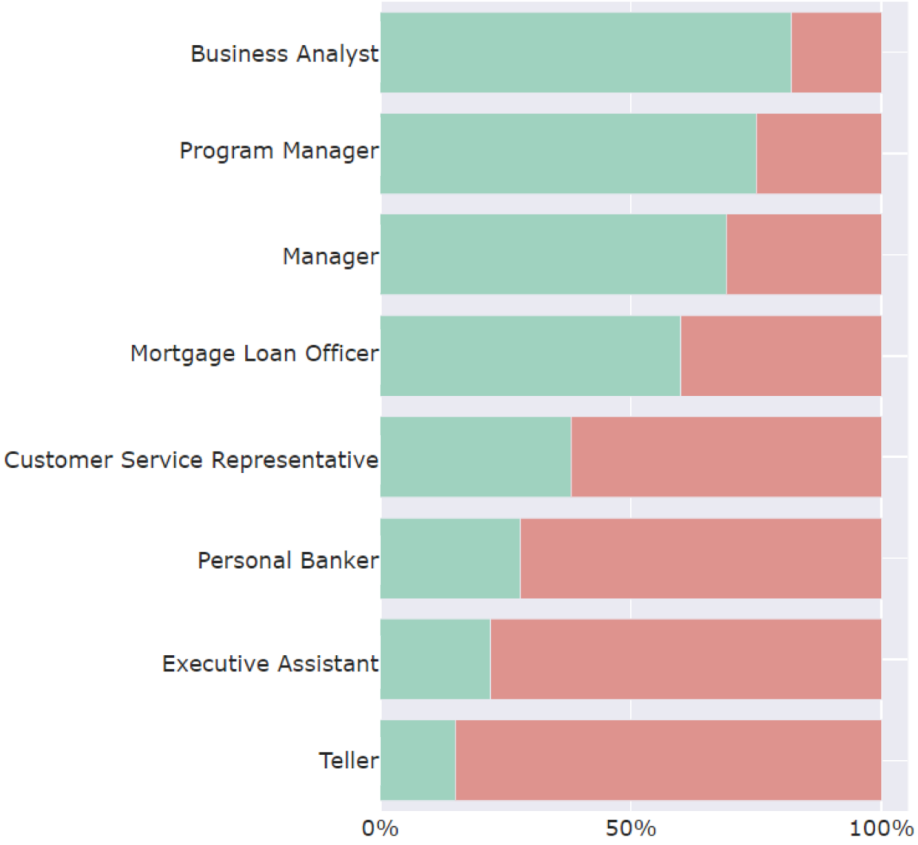
Regions vary in ML exposure



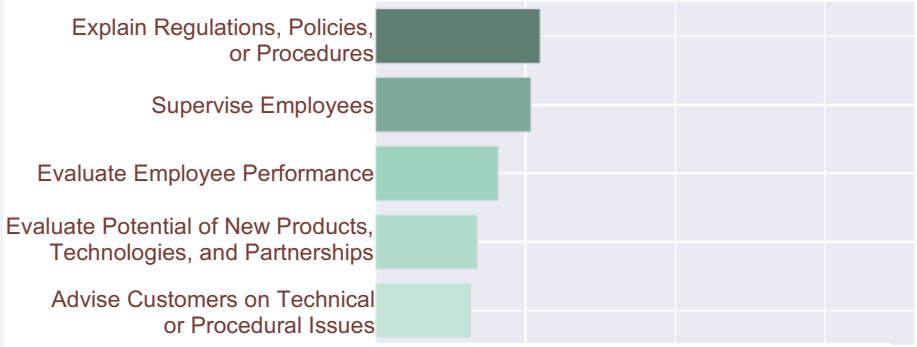
SML Standardized Score by Metro Area

Company: [redacted] Role: All ML/Scan™ Download Data

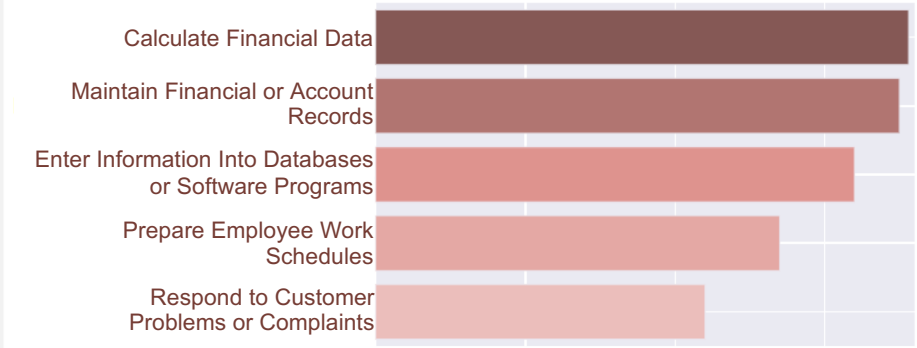
TOP ROLES



LOW EXPOSURE ACTIVITIES

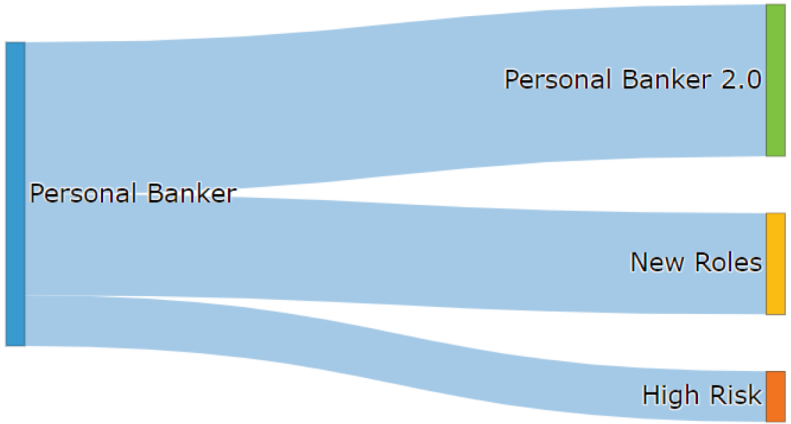


HIGH EXPOSURE ACTIVITIES



Company: [redacted] Role: All ML/Scan™ Download Data

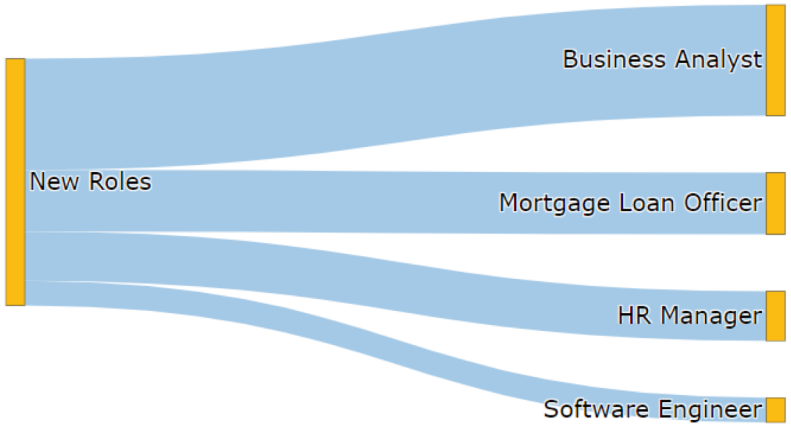
ROLE EVOLUTION



SKILLS OF PERSONAL BANKER 2.0

Low SML Skills	High SML Skills
-Data Analysis	-Cost and Price Calculation
-Managerial Communication	-Data Entry and Categorization
-Leadership	-Credit Authorization
-Product Development and Design	-Customer Record Maintenance
-Customer Relationship Management	-Scheduling

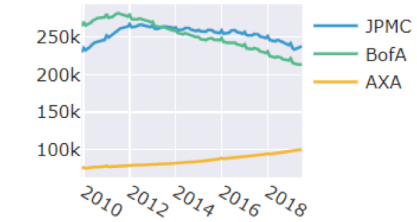
ROLE TRANSITIONS



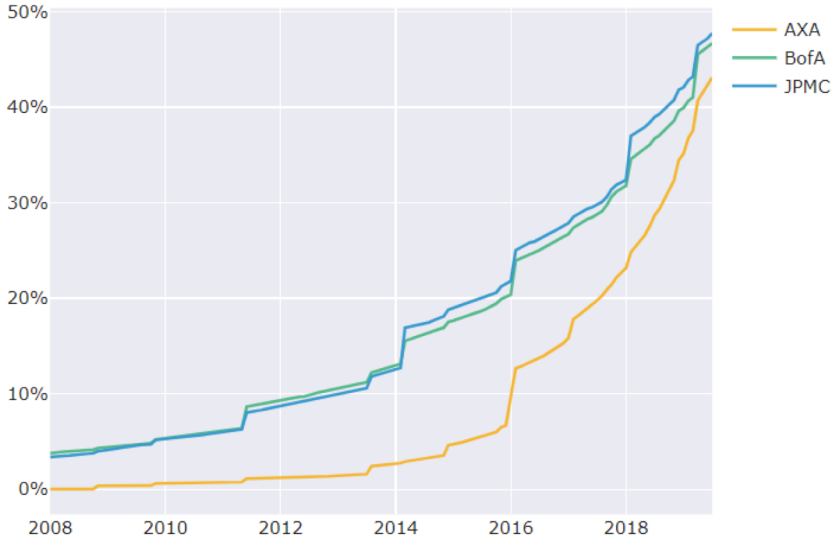
SKILL GAPS FOR TOP TRANSITIONS

Companies: [redacted] Role: All ML/Scan™ Download Data

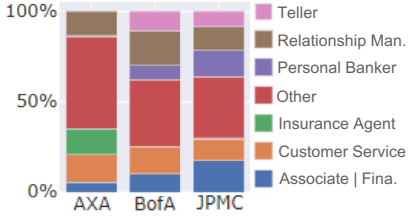
HEADCOUNT



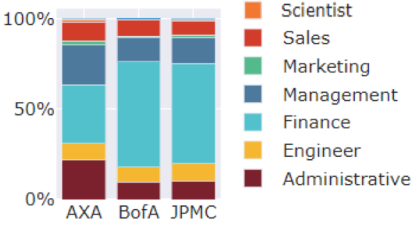
SHARE EXPOSED



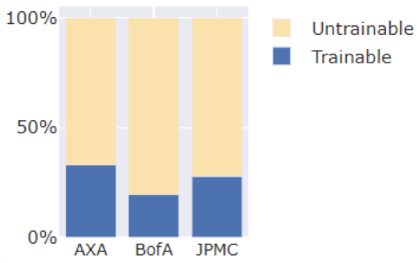
EXPOSED ROLES



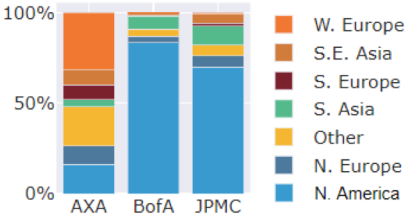
ROLE COMPOSITION



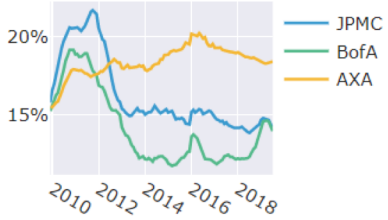
SKILL AGGREGATE



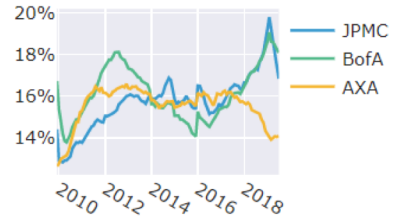
REGION COMPOSITION



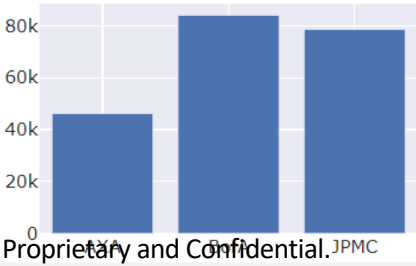
HIRING RATE



ATTRITION RATE



SALARIES



Big data is transforming work…
…and the way we measure work

1. Suitability for Machine Learning (SML) can be assessed via our rubric
   - Can aggregate tasks by occupation, geography, firms and industry
2. ML differs from earlier types of automation (even digital)
   - We cannot simply extrapolate past trends
     - Few occupations will be fully automated
     - Few are immune

   => *Reorganization of jobs will be required to unleash ML*
3. We can create a unique roadmap for each company, city and nation using these data

To learn more, visit http://BrynJolfsson.com/research

ILLUSTRATION: DAVIDE BONAZZI/@SALZMANART