“Work balance analysis in the BigTech companies”

Group H:
Stefania Sferragatta,
Eugenio Baldo,
Clara Lecce

May 2021

1 Introduction

Our project is based on the result provided in the previous work regarding a systematic review related to the human capital in the technology field.
In order to go deeper in our analysis, we chose a dataset which allows us to go further on the initial topic and to focus our interest on which are the factors influencing the employees' wellness.
Before starting the main analysis with the dataset, we present a summary of the most pertinent retrieved papers whose main topic is related to the one of our interest.
After this, as fist step, we display a possible ontology of the topic and, as second one, we explore the chosen dataset in order to begin the analysis focusing on each BigTech company stored.
In the end, we compare the obtained results in order to draw our conclusions.

2 Papers' Analysis

About our Systematic Review, we decided to read some papers resulting from the analysis carried out in the latter, which is the value of Human Capital in Hi-Tech industry.
We selected around nine papers to see what kind of studies were taken in these, and to construct a link from the Systematic Review to our project work, which is about the work balance analysis in the BigTech companies.
All the studies analyzed in the papers take place in the Software Industries, and are about how the Human Resource really interfere with these firms.

2.1 Some information on the papers read

In the first study read [1], they used public data on LinkedIn.com, collecting the CV-like public profiles of IT professionals working in public software firms in US, to make a fim-level analysis and to examine the effects of the human resource inflows and outflows on software firm performance.
They made around six hypothesis about how the human resource inflow and outflow are respectively positively or negatively associated with the firm performance, and in their conclusion (only three out of six hypothesis are supported, expecially the ones about the IT experience and the education level of the human resource inflow) they say that HR managers should consider the different impacts of both the human resource inflow and outflow, in order to develop more effective HR strategies.

Another study [2], investigates the organizational investment in the development of human capital in the context of Open Source Software. They used survey data collected from 114 senior IT managers and IT professionals.
From the hypothesis and consequent analysis they have made, they observed that when firm-specificity is low, organizations may still invest in internal human capital if learning-related scale is justifiably high.

The purpose of another study [3] is to explore the changing role if human resource management in an era of digital transformation. For their research, they used a qualitative research methodology using semi-structured interviews with five HR professionals in the United Arab Emirates. The affected areas were HR planning, reward management, performance management, training and development, health and safety and employee relations and their change in an era of digital transformation.
In this study, the participants tends to focus on how technology has made HR practices and procedures more efficient and effective. In conjunction, companies seem to be using technology more to ease, speed up and improve their current human resource practices and procedures and less to analyse data and plan around such analysis.

There is also an editorial [4] about our topic, that comes to a conclusion where in high-tech companies, a majority of the human capital is highly skilled and educated.

A different study [5] adopts the human capital perspective as a theoretical lens to understand the factors that influence organizational intention to adopt open source software. For their study, they analyze survey responses from 81 chief information officers or information system managers, to analyze how the availability of internal human capital and accessibility to external human capital affect an organization's intention to adopt OSS either directly or indirectly through switching cost.
This study has both theoretical and practical implications. Theoretically, the unique approach from the human capital perspective, influences the innovation on organizational adoption and brings more attention to understanding the influence of human capital in an increasingly technologically environment. Practically, their findings suggest that potential OSS adopters and proponents should develop the human capital

necessary for effective exploitation of OSS.

There is a study [6] about how it is important and critical for many organizations, developing and retaining “IT human capital”.

Another study that we read [7], examines compensation-tenure profiles using salary data collected on IT professionals. Their findings show that compensation increases with organizational tenure for all IT jobs. However, individuals in IT jobs requiring more firm-specific human capital are paid more than those requiring less-firm specific human capital.
Their results suggest the importance of job type in examining compensation-tenure profile in IT, where is interested the human capital, more specifically, the value of the technical one can diminish, while the value of its firm-specific can appreciate with organizational tenure.

In another paper [8], the main question is about how information technology may complement a key, like human capital, as other firm resources. And the answer is mainly crucial to small businesses, because they face a growing demand of IT usage, that they cannot satisfy, because of their limited resources.
In this study, they focus on the analysis of small businesses, and they find that the use of IT services is positively related to firm productivity, which is associated with a combination of human and technology resources.

In the last study read [9], they focus on how human capital investments directed toward employee training are effective in improving employee performance.

3 Work balance analysis of the BigTech companies' employees in California

The aim of our analysis is to focus on which factors have an influence on the work balance equilibrium of an employee in six BigTech companies.

3.1 Ontology of the topic

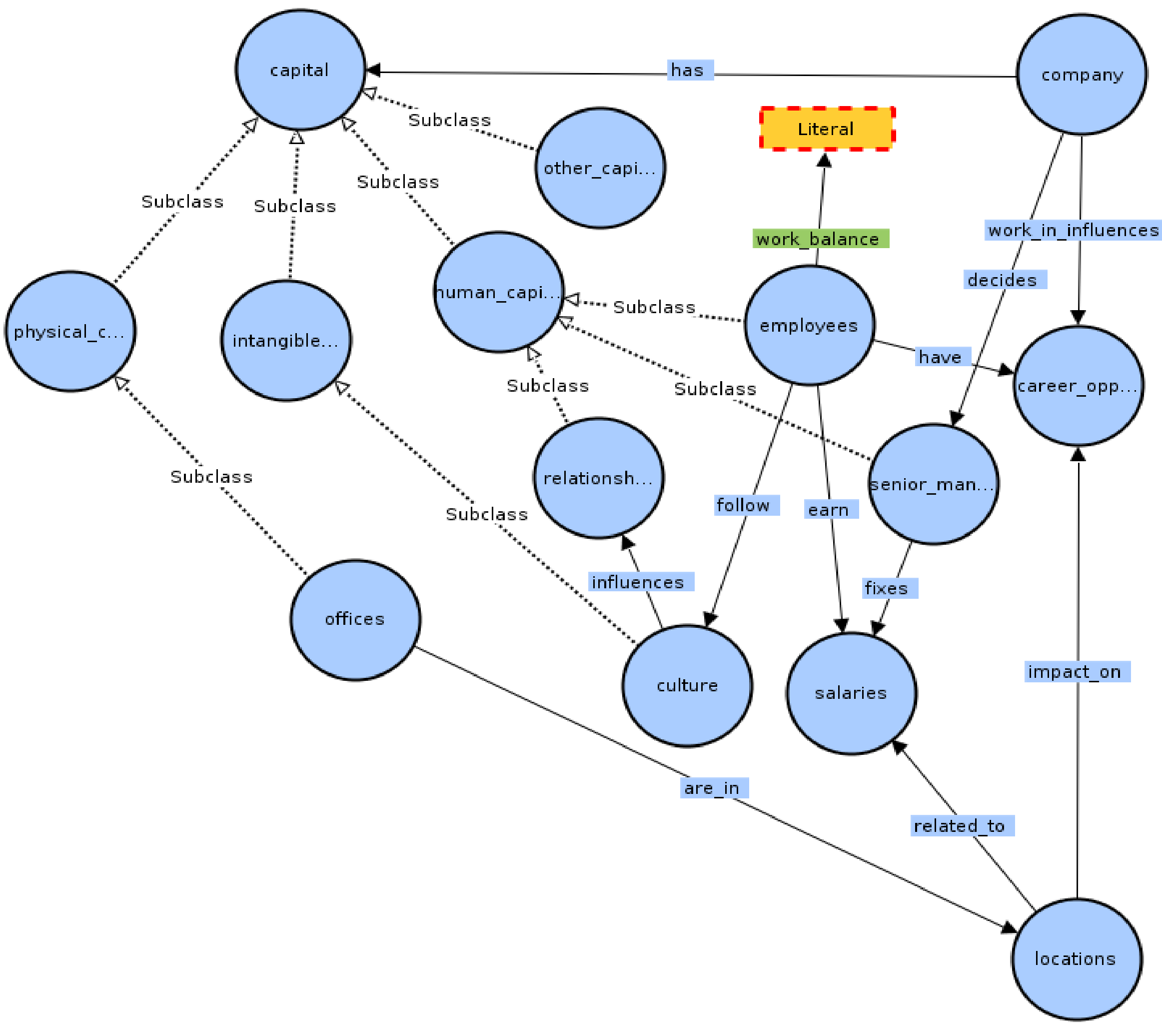

Figure 1: Ontology graph using VOWL on Protègè

3.2 The dataset

The dataset chosen for this analysis was sourced from Glassdoor[1], a website where current and former employees anonymously review companies and their management. It is composed by the following fields:

General variables:

1. Company: Company name;
2. Location : work-place
3. Job-Title: type of job;
4. Year: indicates when the review has been done;
5. Overall Rating: 1–5;

Specific rating variables:

1. Work/Life Balance Rating: 1–5;
2. Culture and Values Rating: 1–5;
3. Career Opportunities Rating: 1–5;
4. Comp Benefits Rating: 1–5;
5. Senior Management Rating: 1–5;

3.3 Textual and visual description of the data

The analysis we carried out is based mainly on the ratings done by the employee that work or have been work in the company. The variables in which we focus our interest are: **work_balance_stars**; **culture_values_stars**; **carrer_opportunities_stars**; **comp_benefit_stars (salary and benefit)**; **senior_management_stars**.
These attributes represent a judgement made by a particular employee between 2008 and 2018 on his job situation in the company on a scale of 1-5.
We chose to focus only on the company located in California (USA) since in that place there are the headquarters of the companies taken into account.
With this restriction we aimed to have less biased observations, with reference to different customs and habits of different regions around the world. Provided that these judgements are likely to be related each other, we would discover these patterns and, in particular, the ones related to the work balance. An hint of the relationships are provided by correlation.

[1]https://www.glassdoor.it/index.htm

3.4 Data Visualization

Below there are the boxplots and the distribution of the ratings with reference to each company.

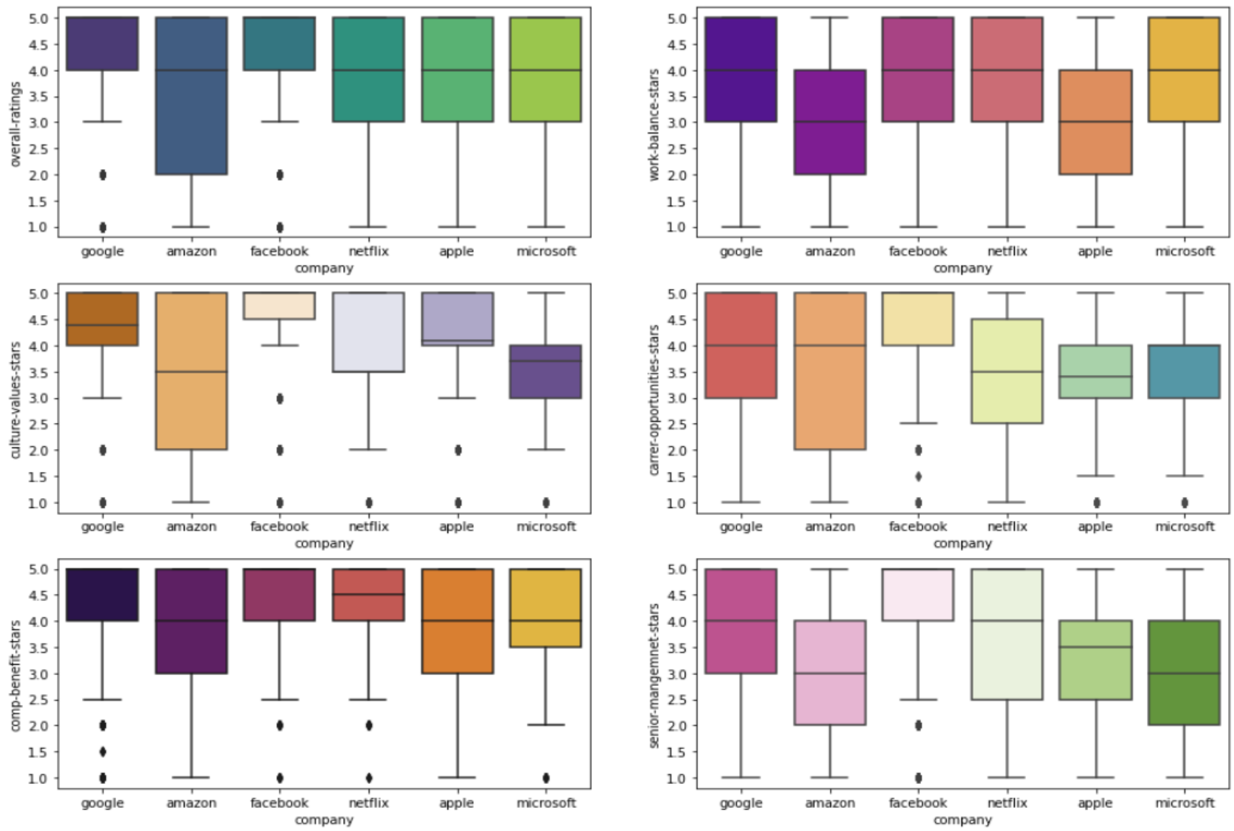

Figure 2: Ratings distributions

By means of the following plots we can find quickly an answer to questions like:

- Which company has better Career Opportunities?
- Which company offers better Work-Life Balance?
- Which company offers better benefits?
- Which company has better Culture Values?

(a) Career opportunities

(b) Work balance

(c) Company's benefit

(d) Company's culture value

Figure 3: Comparison between companies' ratings

3.5 General overview of the analysis

After we analyzed the dataset we focused our attention on company-specific information, hence we continued the analysis on each distinct company in order to obtain meaningful results that at the end we compared to discover whether there are similarities or not among the causal relationships between the companies.
Our goal is to find out which are the factors that may have an influence on the work balance equilibra of an employee and to achieve it we applied for each company in the input dataset some main steps:

1. Analyze the **correlations** among the different variables;
2. Build the **bayesian network** adopting **Hill Climb** search algorithm (which aims to optimize K2 score);
3. Identify the treatment variable and the outcome variable based on the bayesian network generated;
4. Compute inference on the data using as estimator the **Maximum Likelihood** one and as method the **Variable elimination**;

5. Estimate the effect of the treatment on the outcome using **causal inference with a back-door criterion**;

6. **Test the hypothesis** with these tests:"random common cause", "data subset refuter", "placebo treatment refuter";

7. Accept or reject the hypothesis according the **p-value**.

At the end of each analysis we compare the results obtained and use them to make conclusions.

3.6 Description of the methods chosen

3.6.1 Correlation

For our analysis sample, we use the **Pearson correlation coefficient**.
The correlation is a method used to measures how strong is the relationship between two variables. Correlation coeffienct (ρ) is between -1 and 1.
A $\rho = -1$ indicates that the two variables are perfectly negatively correlated, hence the two variables are not correlated and so independent. A $\rho = 1$ indicates that the two variables are perfectly correlated.[10]

3.6.2 Probabilistic Graphical models

The Probabilistic Graphical models (PGM) are statistical models useful for encoding complex relationships thought graphs.[11]
They represent in a compact and visual way joint multivariate distributions and are powerful tool to capture conditional independence relationships between random variables that interact with each other. We used this tool to make inference on the observed data.

3.6.3 Bayesian Network

For our purpose we chose the directed PGM, hence we used for our analysis the Bayesian Network that is a Directed Acyclical Graphs (DAG).
The structure of this model determines the density factorization where the vertices represent random variables and the edges represent the direction in which one variable has influence on (or cause) the other. Bayesian networks have the property that a variable is independent from the non-descendants. [12]

3.6.4 Causal Inference

In many fields, such as economics and public policies, we are not only interested in the data generating process but also in **causal relationships among random variables**. Provided a random vector X composed by random variables (X_1,X_2,...X_N) of interest, with a joint distribution $L(X)$ between them, and assuming there exists a causal graph (DAG) which connects these random variables, we try to infer this causal graph behind

the data generating process.
To obtain the causal graph we start from observations of the joint distribution $L(X)$ and from them some statistical tests are adopted in order to assess the conditional dependencies and in dependencies among the variables.[13]

3.7 Facebook work-balance analysis

The source code for this analysis can be consulted here.

3.7.1 Correlation matrix

	overall_ratings	**work_balance_stars**	**culture_values_stars**	**carrer_opportunities_stars**	**comp_benefit_stars**	**senior_mangemnet_stars**
overall_ratings	1.000000	0.584071	0.767267	0.746212	0.438975	0.750691
work_balance_stars	0.5	1.000000	0.552184	0.496728	0.356884	0.593187
culture_values_stars	0.767267	0.552184	1.000000	0.655536	0.381164	0.720621
carrer_opportunities_stars	0.746212	0.496728	0.655536	1.000000	0.522372	0.680094
comp_benefit_stars	0.438975	0.356884	0.381164	0.522372	1.000000	0.467325
senior_mangemnet_stars	0.750691	0.593187	0.720621	0.680094	0.467325	1.000000

Table 1: Facebook correlation matrix

3.7.2 Bayesian network

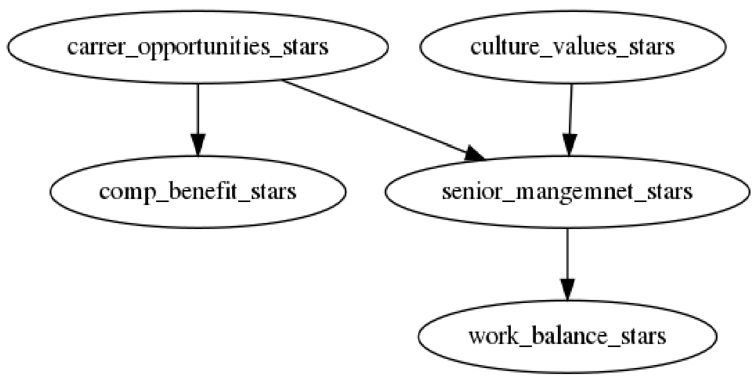

Figure 4: Facebook Bayesian network generated by Hill Climb search algorithm

3.7.3 Conditional probability of the outcome given the treatment

Senior management (treatment)	**Work Balance (Outcome)**	**P(work_balance=i|Senior_management=j)**
5	1	0.0059
5	2	0.0294
5	3	0.1412
5	4	0.2824
5	5	0.5412

Table 2: $P(\text{Outcome}|\text{Treatment})$ for Facebook

3.7.4 Results

Estimated effect and hypothesis testing

- Estimated mean: 0.6494
- p-value: $4\exp{-23}$
- 95% confidence interval: [0.5311,0.7676]
- Hypothesis testing with random common cause:
  estimated effect: 0.6494, new effect: 0.6519
- Hypothesis testing with subset of data: estimated effect: 0.6494, new effect: 0.6495, p-value: 0.4200
- Hypothesis testing with placebo treatment: estimated effect: 0.6494, new effect: 0.0076,p-value: 0.46

3.8 Netflix work-balance analysis

The source code for this analysis can be consulted here.

3.8.1 Correlation matrix

	overall_ratings	work_balance_stars	culture_values_stars	carrer_opportunities_stars	comp_benefit_stars	senior_mangemnet_stars
overall_ratings	1.000000	0.680999	0.626473	0.800626	0.517923	0.810433
work_balance_stars	0.680999	1.000000	0.463042	0.550973	0.310474	0.682019
culture_values_stars	0.626473	0.463042	1.000000	0.571339	0.405328	0.603752
carrer_opportunities_stars	0.800626	0.550973	0.571339	1.000000	0.464276	0.718294
comp_benefit_stars	0.517923	0.310474	0.405328	0.464276	1.000000	0.433990
senior_mangemnet_stars	0.810433	0.682019	0.603752	0.718294	0.433990	1.000000

Table 3: Netflix correlation matrix

3.8.2 Bayesian network

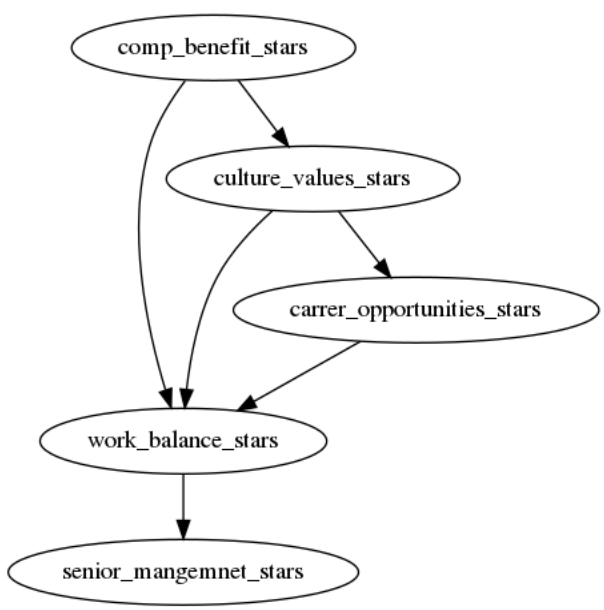

Figure 5: Netflix Bayesian network generated by Hill Climb search algorithm

3.8.3 Conditional probability of the outcome given the treatment

W_B(Outcome)	P(W_B=i|C_B)	P(W_B=i|C_V)	P(W_B=i|C_O)
1	0.0525	0.0056	0.0125
2	0.1111	0.0056	0.0564
3	0.1895	0.1792	0.1144
4	0.2831	0.3181	0.4392
5	0.3639	0.4917	0.3775

Table 4: $P(\text{Outcome}|\text{Treatment}{=}5)$ for Netflix

3.8.4 Results

Estimated effect and hypothesis testing on the distinct treatments

- Estimated mean: (career opp,0.2578),(culture values, 0.4852),(compensation, 0.5398)
- p-value: (career opp.,0.0149),(culture values, $2.13\exp^{-6}$),(compensation, 0.0004)
- 95% confidence interval:(career opp.,[0.0514,0.4642]), (culture values,[0.2946,0.6759]), (compensation,[0.2483,0.8312])
- Hypothesis testing with random common cause:
  estimated effect: (career opp. 0.2578),(culture values, 0.4852),(compensation,0.5398)
  new effect: (career opp. 0.2600),(culture values, 0.4834),(compensation,0.5467)
- Hypothesis testing with subset of data:
  estimated effect: (career opp,0.2578),(culture values, 0.4852),(compensation, 0.5398)
  new effect: (career opp, 0.2773),(culture values 0.4913),(compensation, 0.5497)
  p-value: (career opp,0.37),(culture values, 0.42),(compensation 0.48)
- Hypothesis testing with placebo treatment: estimated effect: (career opp.,0.2578),(culture values, 0.4852),(compensation, 0.5397) new effect: (career opp. 0.0025),(culture values, 0.0099),(compensation,-0.0003) p-value:(career opp. 0.4599),(culture values, 0.47),(compensation, 0.49)

3.9 Microsoft work-balance analysis

The source code for this analysis can be consulted here.

3.9.1 Correlation matrix

	overall_ratings	work_balance_stars	culture_values_stars	carrer_opportunities_stars	comp_benefit_stars	senior_mangemnet_stars
overall_ratings	1.000000	0.507352	0.635737	0.649684	0.507082	0.702221
work_balance_stars	0.507352	1.00000	0.418678	0.389408	0.406673	0.424974
culture_values_stars	0.635737	0.418678	1.000000	0.454501	0.344502	0.615337
carrer_opportunities_stars	0.649684	0.389408	0.454501	1.000000	0.495858	0.574876
comp_benefit_stars	0.507082	0.406673	0.344502	0.495858	1.000000	0.359791
senior_mangemnet_stars	0.702221	0.424974	0.615337	0.574876	0.359791	1.000000

Table 5: Microsoft correlation matrix

3.9.2 Bayesian network

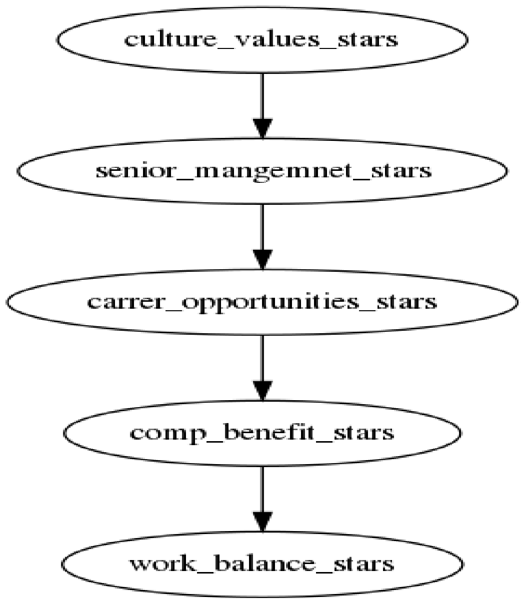

Figure 6: Microsoft Bayesian network generated by Hill Climb search algorithm

3.9.3 Conditional probability of the outcome given the treatment

W_B (Outcome)	P(W_B=i|S_M=j)
1	0.0000
2	0.1087
3	0.1739
4	0.2391
5	0.4783

Table 6: $P(\text{Outcome}|\text{Treatment})$ for Microsoft

3.9.4 Results

Estimated effect and hypothesis testing

- estimated mean: 0.4113
- p-value: $3.54\,\exp^{-5}$
- 95% confidence interval: [0.2206,0.6020]
- hypothesis testing with random common cause: estimated effect: 0.4113, new effect: 0.4097
- hypothesis testing with subset of data: estimated effect: 0.4113, new effect: 0.4116, p-value: 0.4200
- hypothesis testing with placebo treatment: estimated effect: 0.4113, new effect: 0.0004, p-value: 0.47

3.10 Amazon work-balance analysis

The source code for this analysis can be consulted here.

3.10.1 Correlation matrix

	overall_ratings	**work_balance_stars**	**culture_values_stars**	**carrer_opportunities_stars**	**comp_benefit_stars**	**senior_mangemnet_stars**
overall_ratings	1.000000	0.658603	0.755126	0.721668	0.543276	0.735683
work_balance_stars	0.658603	1.000000	0.621683	0.510959	0.431752	0.625463
culture_values_stars	0.755126	0.621683	1.000000	0.646568	0.474040	0.726017
carrer_opportunities_stars	0.721668	0.510959	0.646568	1.000000	0.550387	0.673629
comp_benefit_stars	0.543276	0.431752	0.474040	0.550387	1.000000	0.498529
senior_mangemnet_stars	0.735683	0.625463	0.726017	0.673629	0.498529	1.000000

Table 7: Amazon correlation matrix

3.10.2 Bayesian network

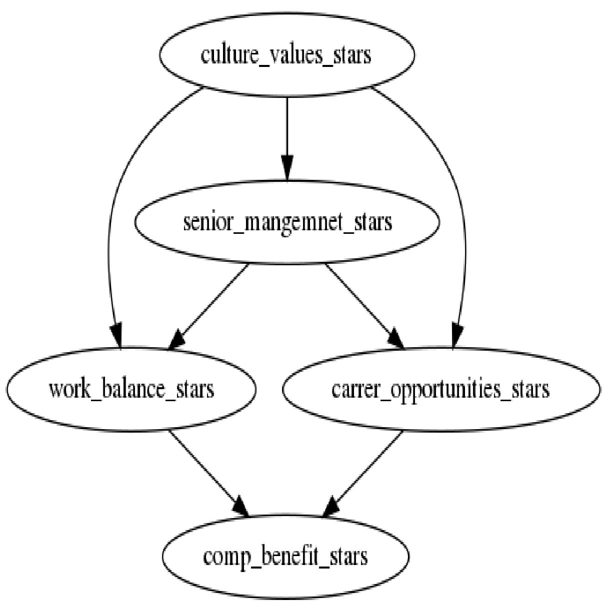

Figure 7: Amazon Bayesian network generated by Hill Climb search algorithm

3.10.3 Conditional probability of the outcome given the treatment

W_B (Outcome)	P(W_B=i|S_M=j)	P(W_B=i|C_V=j)
1	0.0211	0.0309
2	0.0411	0.0702
3	0.1522	0.1966
4	0.2967	0.3085
5	0.4890	0.3938

Table 8: $P(\text{Outcome}|\text{Treatment=5})$ for Amazon

3.10.4 Results

Estimated effect and hypothesis testing

- Estimated mean: (culture values, 0.6075), (senior mangment, 0.3701)
- p-value: (culture values, ≈ 0.0), (senior mangment, $1.67127829 \exp^{-128}$)
- 95% confidence interval:(culture values,[0.5864561,0.62867036]), (senior mangment,[0.34089128, 0.39933306])
- Hypothesis testing with **random common cause**:
  Estimated effect: (culture values, 0.6075), (senior mangment, 0.3701)
  New effect: (culture values, 0.6075), (senior mangment, 0.3701)
- Hypothesis testing with **subset of data**:
  Estimated effect: (culture values, 0.6075), (senior mangment, 0.3701)
  New effect: (culture values 0.6072), (senior mangment, 0.3713)
  p-value: (culture values, 0.48), (senior mangment 0.49)
- Hypothesis testing with **placebo treatment**:
  Estimated effect: (culture values, 0.6075), (senior mangment,0.3701)
  New effect: (culture values, 0.0032), (senior mangment,-0.0006)
  p-value: (culture values, 0.45), (senior mangment, 0.46)

3.11 Apple work-balance analysis

The source code for this analysis can be consulted here.

3.11.1 Correlation matrix

	overall_ratings	work_balance_stars	culture_values_stars	carrer_opportunities_stars	comp_benefit_stars	senior_mangemnet_stars
overall_ratings	1.000000	0.562918	0.652562	0.653245	0.509659	0.684976
work_balance_stars	0.562918	1.000000	0.476680	0.443474	0.369437	0.543645
culture_values_stars	0.652562	0.476680	1.000000	0.487426	0.393125	0.605700
carrer_opportunities_stars	0.653245	0.443474	0.487426	1.000000	0.474867	0.617904
comp_benefit_stars	0.509659	0.369437	0.393125	0.474867	1.000000	0.418552
senior_mangemnet_stars	0.684976	0.543645	0.605700	0.617904	0.418552	1.000000

Table 9: Apple correlation matrix

3.11.2 Bayesian network

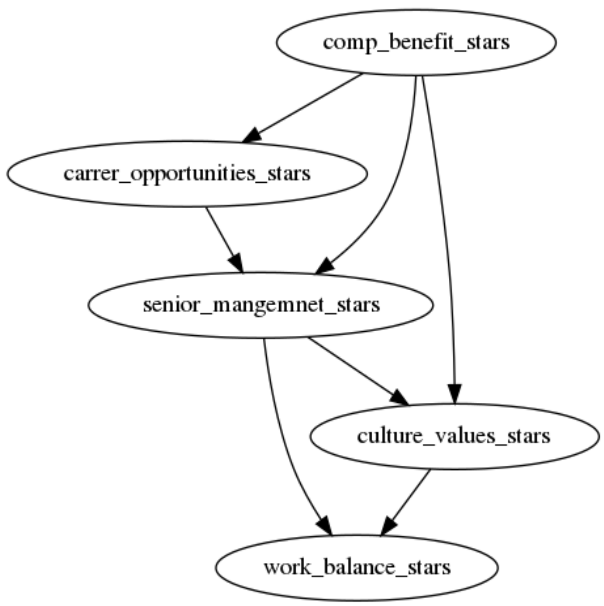

Figure 8: Apple Bayesian network generated by Hill Climb search algorithm

3.11.3 Conditional probability of the outcome given the treatment

W_B (Outcome)	P(W_B=i|S_M=j)	P(W_B=i|C_V=j)
1	0.0172	0.0367
2	0.0613	0.1091
3	0.1513	0.2141
4	0.2318	0.2907
5	0.5383	0.3494

Table 10: $P(\text{Outcome}|\text{Treatment=5})$ for Apple

3.11.4 Results

Estimated effect and hypothesis testing

- Estimated mean: (culture values, 0.2090), (senior mangment, 0.4165)
- p-value: (culture values, 2.83212639 $\exp^{-15}$), (senior mangment, 5.81067644 $\exp^{-78}$)
- 95% confidence interval:(culture values,[0.15747743,0.26054387]), (senior mangment,[0.37456612,0.45863136])
- Hypothesis testing with **random common cause**:
  Estimated effect: (culture values, 0.2090), (senior mangment, 0.4165)
  New effect: (culture values, 0.2095), (senior mangment, 0.4177)
- Hypothesis testing with **subset of data**:
  Estimated effect: (culture values, 0.2090), (senior mangment, 0.4165)
  New effect: (culture values 0.2081), (senior mangment, 0.4171)
  p-value: (culture values, 0.47), (senior mangment 0.5)
- Hypothesis testing with **placebo treatment**:
  Estimated effect: (culture values, 0.2090), (senior mangment, 0.4165)
  New effect: (culture values,-0.0010), (senior mangment,-0.0002)
  p-value: (culture values, 0.5), (senior mangment, 0.43)

3.12 Google work-balance analysis

The source code for this analysis can be consulted here.

3.12.1 Correlation matrix

	overall_ratings	work_balance_stars	culture_values_stars	carrer_opportunities_stars	comp_benefit_stars	senior_mangemnet_stars
overall_ratings	1.000000	0.521767	0.616378	0.641013	0.482167	0.647143
work_balance_stars	0.521767	1.000000	0.497289	0.410909	0.371769	0.518078
culture_values_stars	0.616378	0.497289	1.000000	0.489991	0.411630	0.605130
carrer_opportunities_stars	0.641013	0.410909	0.489991	1.000000	0.477543	0.591095
comp_benefit_stars	0.482167	0.371769	0.411630	0.477543	1.000000	0.453748
senior_mangemnet_stars	0.647143	0.518078	0.605130	0.591095	0.453748	1.000000

Table 11: Google correlation matrix

3.12.2 Conditional probability of the outcome given the treatment

W_B (Outcome)	P(W_B=i|S_M=j)	P(W_B=i|C_V=j)
1	0.0000	0.0034
2	0.0158	0.0307
3	0.0976	0.1056
4	0.1847	0.2453
5	0.7018	0.6150

Table 12: P(Outcome|Treatment=5) for Google

3.12.3 Bayesian network

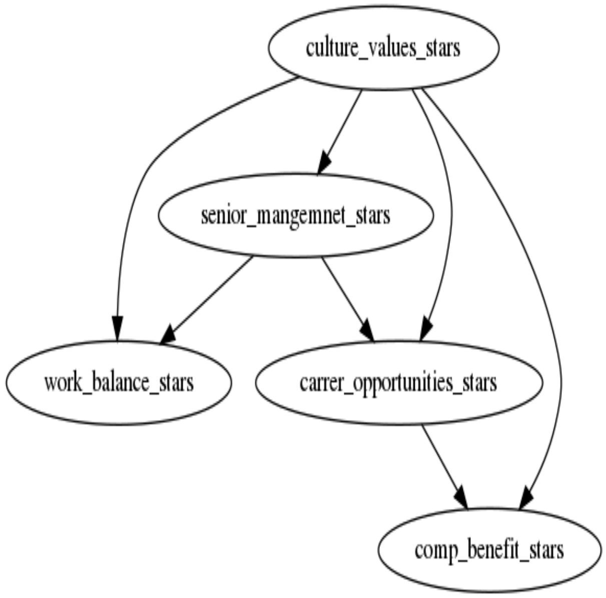

Figure 9: Google Bayesian network generated by Hill Climb search algorithm

3.12.4 Results

Estimated effect and hypothesis testing

- Estimated mean: (culture values, 0.5924), (senior mangment, 0.3263)
- p-value: (culture values, $1.49853556\,\exp^{-81}$), (senior mangment, $1.1502112\,\exp^{-30}$)
- 95% confidence interval:(culture values,[0.53612757,0.64883885]),
  (senior mangment,[0.27225435,0.38053441])
- Hypothesis testing with **random common cause**:
  Estimated effect: (culture values, 0.5924), (senior mangment, 0.3263)
  New effect: (culture values, 0.5921), (senior mangment, 0.3261)
- Hypothesis testing with **subset of data**:
  Estimated effect: (culture values, 0.5924), (senior mangment, 0.3263)
  New effect: (culture values0.5922), (senior mangment, 0.3284)
  p-value: (culture values, 0.47), (senior mangment 0.41)
- Hypothesis testing with **placebo treatment**:
  Estimated effect: (culture values,0.5924), (senior mangment,0.3263)
  New effect: (culture values,0.5924), (senior mangment,0.0029)
  p-value: (culture values, 0.49), (senior mangment, 0.42)

3.13 Results and comparison

Company	**Treatment**	**Average Effect**	**95% C.I**	**Hypothesis accepted**
Facebook	SM	0.6494	[0.5311,0.7676]	Yes
Netflix	CO	0.2578	[0.0514,0.4642]	Yes
Netflix	CV	0.4852	[0.2946,0.6759]	Yes
Netflix	Comp	0.5398	[0.2483,0.8312]	Yes
Microsoft	Comp	0.4113	[0.2206,0.6020]	Yes
Amazon	CV	0.6075	[0.5865,0.6287]	Yes
Amazon	SM	0.3701	[0.3408, 0.3993]	Yes
Apple	CV	0.2090	[0.1575, 0.2605]	Yes
Apple	SM	0.4165	[0.3746, 0.4586]	Yes
Google	CV	0.5924	[0.5361, 0.6488]	Yes
Google	SM	0.3263	[0.2725, 0.3805]	Yes

Table 13: Summary table of the results obtained

Legend for reading the abbreviations of the Treatments:

- SM = Senior Managment,
- CO = Career opportunities,
- CV = Culture Values,
- Comp = Comp benefits.

3.14 Conclusions

The goal of our project was to discover which factors may influence on the work balance of the employees of the big tech companies, with reference to California.
To do this, we applied the causal analysis through these main steps:

1. Model a causal inference problem using an assumption;
2. Identify an expression for the causal effect under this assumption;
3. Estimate the expression using statistical methods;
4. Verify the validity of the estimate using a variety of robustness checks.

From our results obtained we can deduct that :

- Senior Management and Culture values are the variables which impact more on the work balance overall.
- Companies like Amazon, Apple and Google have the same treatments on the work balance. Especially Google and Amazon have very similar results on the effect on the outcome variables. This may imply that the two organizations are structured in a comparable way.
- Compensation incidence is significant on Netflix and Microsoft, although the confidence intervals are wide.
- Overall, the treatments found with Hill Climb Search tend to be the most correlated with the work balance.
- The breadth of the confidence intervals for Netflix and Microsoft gives some doubts about the consistency of the results for these companies.

In this scenario the Hill Climb search, which is an heuristic algorithm, results as very effective given that all the treatment outcome hypothesis have been accepted.

3.15 Further analysis

It is possible to widen our analysis not only to the ratings variables, which are structured data, but also considering the review text, which is an unstructured data type. Using text field, we may discover additional relationships and common patterns, that may confirm or not our previous results and beliefs as well as giving additional information and insights related to employees' wellness in the Big tech companies.

References

[1] Ge, Chunmian and Zou, Xiao, **"Human Resource Flow within Software Industry: A Firm-Level Investigation"** (2013). *PACIS 2013 Proceedings.*

[2] LI, Y., TAN, C.H. and TEO, H.H. (2008). **Firm-Specificity and Organizational Learning-related Scale on In-vestment in Internal Human Capital for Open Source Software Adoption.** In: *2008 SIGMIS Computer Personnel Doctoral Consortium and Research Conference Proceedings.* Association for Computing Machinery (ACM).

[3] **THE CHANGING ROLE OF HUMAN RESOURCE MANAGEMENT IN AN ERA OF DIGITAL TRANSFOR-MATION.** Fenech, Roberta; Baguant, Priya; Ivanov, Dan. *Journal of Management Information and Decision Sciences; Weaverville* Vol. 22, Fasc. 2, (2019): 1-10.

[4] Jolly, D.R. (2005) **'Editorial: human resource management in high-tech companies'**, *Int. J. Technology Management*, Vol. 31, Nos. 3/4, pp.197–203.

[5] Li, Yan; Tan, Chuan-Hoo; Teo, Hock-Hai; and Siow, Alex, **"A Human Capital Perspective of Organizational Intention to Adopt Open Source Software"** (2005). *ICIS 2005 Proceedings.* 12.

[6] R. A. Josefek and R. J. Kauffman, **"Five degrees of separation: a human capital model of employment-related decisionmaking in the information technology workforce"**, *Proceedings of the 32nd Annual Hawaii Interna-tional Conference on Systems Sciences.* 1999. HICSS-32. Abstracts and CD-ROM of Full Papers, 1999, pp. 13 pp.-, doi: 10.1109/HICSS.1999.772695.

[7] Sandra Slaughter and Soon Ang. 2003. **Returns to human capital for information technology professionals.** *In Proceedings of the 2003 SIGMIS conference on Computer personnel research: Freedom in Philadelphia–leveraging differences and diversity in the IT workforce (SIGMIS CPR '03).* Association for Computing Ma-chinery, New York, NY, USA, 143–146.

[8] Ana Cristina O. Siqueira Maria Tereza L. Fleury (2011) **Complementarities of human capital and infor-mation technology: small businesses, emerging economy context and the strategic role of firm resources**, *Technology Analysis Strategic Management*, 23:6, 639-653, DOI: 10.1080/09537325.2011.585032

[9] Ravi Bapna, Nishtha Langer, Amit Mehra, Ram Gopal, Alok Gupta, **Human Capital Investments and Employee Performance: An Analysis of IT Services Industry.** *Management Science 59* (3) 641-658 https://doi.org/10.1287/mnsc.1120.1586

[10] Benesty J., Chen J., Huang Y., Cohen I. (2009) Pearson Correlation Coefficient. In: Noise Reduction in Speech Processing. Springer Topics in Signal Processing, vol 2. Springer, Berlin, Heidelberg

[11] Probabilistic graphical models: principles and techniques - Koller, D.; Friedman, N, MIT Press, 2009.

[12] Serena H. Chen, Carmel A. Pollino, Good practice in Bayesian network modelling, Environmental Modelling Software, Volume 37, 2012, Pages 134-145, ISSN 1364-8152

[13] Pearl, J. (2000, 2009), Causality. Models, Reasoning, and Inference. Cambridge University Press.

[14] DoWhy library documentation for causal inference in python

[15] Pgmpy library documentation for probabilistic graphical models in python

[16] Protègè software for ontologies

[17] Seminar on ontologies with Protègè and Eddy, S. di Leo, 15/04/2021, Sapienza University, Rome

[18] Seminar on Causal Inference, A. Moneta, 26/04/2021, Sapienza University, Rome

[19] Seminar on probabilistic graphical models, G. Quaglia, 12/04/2021, Sapienza University, Rome