

PyCon 2019

Ever wondered which company best suits you?

Myself

- Education
 - BITS Pilani (B.E. in Civil Engineering)
 - ESSEC Business School (Master in Data Sciences and Business Analytics)
- Work Experience
 - TESCO
 - eBay
- Domain Knowledge
 - Banking
 - Promotions and Coupons
 - Pricing (delivery charges & commission on sale of items)
 - Incremental Analysis
 - A/B Testing


How do I find the best company ?

- Define “Best”
 - Salary ?
 - Work-Life balance ?
 - Management and Manager ?
 - Career Growth ?
 - Campus and Cafeteria ?
 - Promotions and Hikes ?
 - State of Technology used ?
 - Transport ?

Few numbers can answer.

Text would give the nitty-gritty details.

Leverage from Glassdoor website



Amazon

[Overview](#)[35k Reviews](#)[32k Jobs](#)[45k Salaries](#)[21k Interviews](#)[11k Benefits](#)[551 Photos](#)

Amazon Reviews

Search job titles

Find Reviews

Clear AllFull-time, Part-time X

Filter

Found 4822 reviews

Sort: Popular

4.2

4 stars


Rating Trends

84%

Recommend to a Friend


95%

Approve of CEO



Jeff Bezos3290 Ratings

Aug 8, 2019



"Nice place to work"

★★★★★

Former Employee - Lead AVS in Chennai

Recommend

Positive Outlook

Approves of CEO

I worked at Amazon full-time for more than a year

Pros

Very good leave policy Work from home Great work environment

Cons

Growth in the organization is bit slow

f

t


e

l

Helpful

Aug 5, 2019

Helpful (1)



"extremely good."

★★★★★


Current Employee - Senior Compliance Associate in Bangalore

Recommend

Positive Outlook

Approves of CEO

I have been working at Amazon full-time for less than a year



"Nice place to work"

★★★★★

Former Employee - Lead AVS in Chennai

Recommend

Positive Outlook

Approves of CEO

I worked at Amazon full-time for more than a year

Pros

Very good leave policy Work from home Great work environment

Cons

Growth in the organization is bit slow

f

t

e

l

Helpful

Work/Life Balance

Culture & Values

Career Opportunities

Compensation and Benefits

Senior Management

Methodology

1. Web-scrape
2. Insights from numerical data (salary/ratings/work experience)
 - Trends
 - Slice & Dice
3. Insights from text (Pros/Cons/Advice to Management)
 - Form tf-idf vector matrix
 - Dimensional reduction
 - Cluster reviews
 - Identify key words to label topics

Web-Scrape Reviews, Ratings and Salaries

```
import urllib.request, sys
from bs4 import BeautifulSoup
from pandas import DataFrame
```

```
Emp_ID = []
Date = []
Description = []
Overall_Rating = []
Work_Exp = []
Job_Title = []
Author_Loc = []
Work_Life_Rating = []
Culture_Values_Rating = []
Career_Opp_Rating = []
Comp_Benefits_Rating = []
Senior_Mgmt_Rating = []
Pros = []
Cons = []
Advice_Mgmt = []
```

#Following scrapes data from page 1

```
page_url = "https://www.glassdoor.co.in/Reviews/Tesco"
page_hdr = {'User-Agent': 'Chrome'}
page_req = urllib.request.Request(page_url, headers =
page_response = urllib.request.urlopen(page_req)
page_soup = BeautifulSoup(page_response, "html.parser")
```

```
page_containers = page_soup.findAll("li", {"class": "
for container in page_containers:
#-----
#-----
    Emp_ID.append(container['id'].strip())
    try:
        Date.append(str(container.find("time", {"class": "
    except:
```

#Sample for one page

```
Profile = []
Samples = []
Salary_Range = []
Min_Salary = []
Max_Salary = []
Min_Units = []
Max_Units = []
```

```
url = "https://www.glassdoor.co.in/Salary/Tesco-Bengaluru-Salaries-E341690.htm"
```

```
hdr = {'User-Agent': 'Chrome'}
```

```
req = urllib.request.Request(url, headers=hdr)
```

```
response = urllib.request.urlopen(req)
```

```
soup = BeautifulSoup(response, "html.parser")
```

```
containers = soup.findAll("div", {"class": "row"})
```

```
print("start of pg1")
```

```
for container in containers:
```

```
    Profile.append(container.find("div", {"class": "employerSalaryInfo tightVert"}).span.text.strip())
```

```
    Samples.append(container.find("div", {"class": "salaryCount minor"}).text.strip())
```

```
    Salary_Range.append(container.find("div", {"class": "hideDesk meanPay padTop aboveRangeBar"}).strong.text.strip())
```

```
    Min_Units.append(container.find("div", {"class": "minor cell alignLt minPay"}).text.strip()[-1:])
```

```
    Max_Units.append(container.find("div", {"class": "minor cell alignRt maxPay"}).text.strip()[-1:])
```

```
    Min_Salary.append(container.find("div", {"class": "minor cell alignLt minPay"}).text.strip()[1:-1])
```

```
    Max_Salary.append(container.find("div", {"class": "minor cell alignRt maxPay"}).text.strip()[1:-1])
```

```
data = {"Profile": Profile, "Samples": Samples, "Salary_Range": Salary_Range,
        "Min_Salary": Min_Salary, "Max_Salary": Max_Salary, "Min_Units": Min_Units, "Max_Units": Max_Units}
```

```
frame = DataFrame(data, columns= ["Profile", "Samples", "Salary_Range", "Min_Salary", "Max_Salary", "Min_Units", "Max_Units"])
```

```
frame.to_csv('sal1.csv')
```

start of pg1

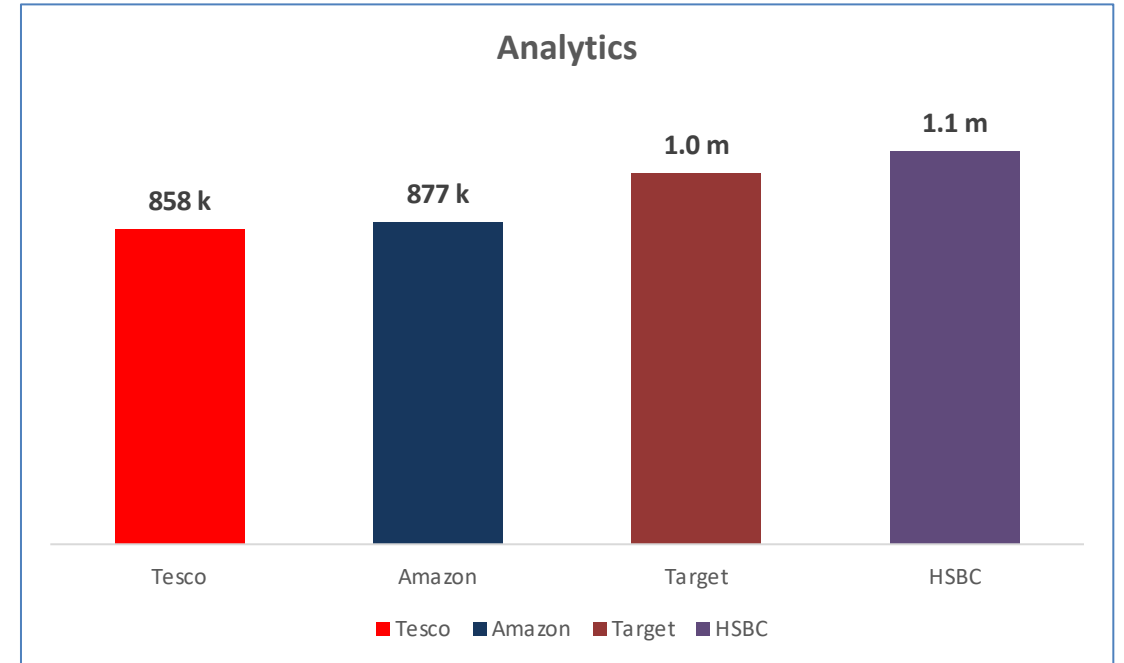
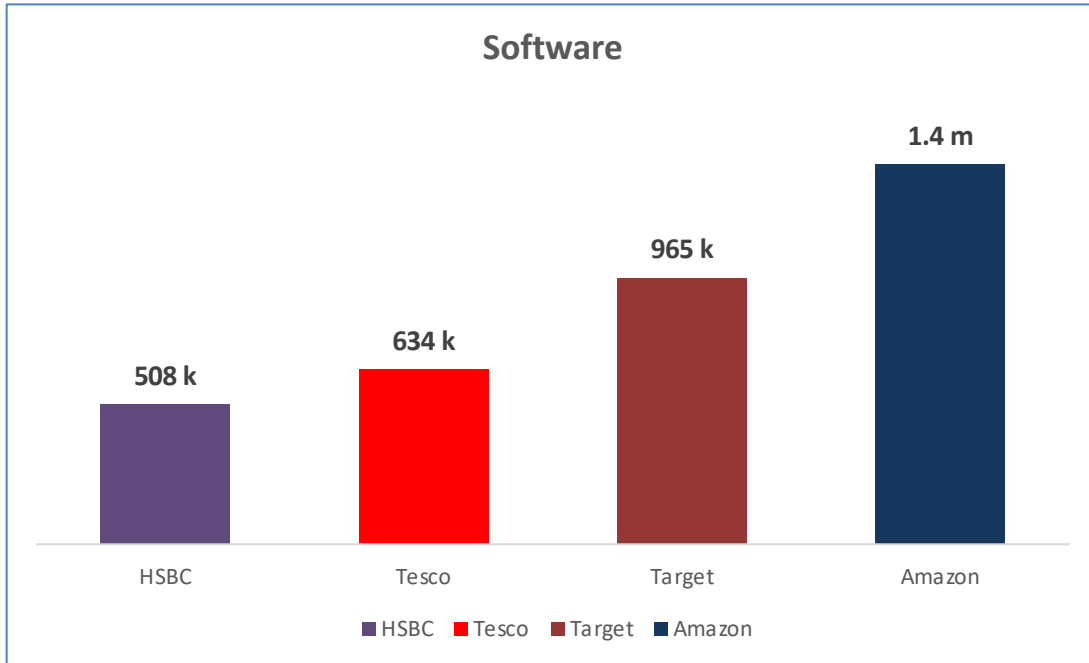
Website Data in Rows and Columns – Reviews & Ratings

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | R |
|----|----|-----------|-----------|-----------------------------|-------------|-----------------------|------------|-------------|------------------------------|---------------------------------------|-------------------|---------------|----------------|---------------|---------------|-------------|---|
| 1 | | Emp_ID | Date | Description | Overall_Rat | Work_Exp | Job_Title | Author_Loc | Pros | Cons | Advice_Mgmt | Work_Life_Rat | Culture_Values | Career_Opp_Ra | Comp_Benefits | Senior_Mgmt | R |
| 2 | 0 | empReview | 15-Oct-17 | "Tesco Bengaluru Review" | 4 | I have been working a | Current Em | NA | One of the best place to w | Company Transport is r | NA | 3 | 4 | 4 | 4 | 4 | |
| 3 | 1 | empReview | 9-Oct-17 | "Very friendly work cultur | 5 | I have been working a | Current Em | Bengaluru | Flexible working hours , fr | Nothing much , but sor | Nothing | 5 | 5 | 4 | 5 | 5 | |
| 4 | 2 | empReview | 16-Oct-17 | "Great environment but l | 3 | I worked at Tesco Ben | Former Em | Bengaluru | The teams interact very clc | Lot of politics at the wo | NA | 4 | 4 | 3 | 3 | 4 | |
| 5 | 3 | empReview | 15-Oct-17 | "Working Experience" | 3 | I worked at Tesco Ben | Former Em | Bengaluru | Work culture in Tesco is ge | There is less fun activiti | NA | 4 | 4 | 3 | 3 | 4 | |
| 6 | 4 | empReview | 8-Oct-17 | "New recruit" | 3 | I have been working a | Current Em | Bengaluru | Good salary, around 17lpc | Bad culture, bad work, NA | NA | 4 | 5 | 3 | 5 | 3 | |
| 7 | 5 | empReview | 22-Sep-17 | "Awesome" | 4 | I worked at Tesco Ben | Former Em | Bengaluru | Nice Green Campus and ha | I was in Tesco for about NA | NA | 5 | 4 | 4 | 4 | 4 | |
| 8 | 6 | empReview | 5-Oct-17 | "Good company to learn i | 3 | I have been working a | Current Em | Bengaluru | Great Manager, Great Tear | Unstable leadership(To | NA | 5 | 4 | 2 | 4 | 1 | |
| 9 | 7 | empReview | 4-Oct-17 | "SDE1" | 3 | I have been working a | Current Em | NA | 1.lot to learn new technol | No such negatives that | NA | 3 | 4 | 3 | 3 | 3 | |
| 10 | 8 | empReview | 18-Sep-17 | "Great Place to work" | 4 | I worked at Tesco Ben | Former Em | Bengaluru | Good work life balance at | Teams and manager cha | NA | 4 | 4 | 4 | 4 | 4 | |
| 11 | 9 | empReview | 2-Oct-17 | "Good company" | 3 | I worked at Tesco Ben | Former Em | Bengaluru | Competitive salaryGreat w | You wont get to work o | NA | 5 | 4 | 3 | 4 | 3 | |
| 12 | 10 | empReview | 1-Oct-17 | "Senior Manager - Faciliti | 3 | I have been working a | Current Em | Bengaluru | Sufficient Flexibility, Empl | Leadership decisions ar | NA | NA | NA | NA | NA | NA | |
| 13 | 11 | empReview | 14-Sep-17 | "Payroll officer experienc | 5 | I worked at Tesco Ben | Former Em | Whitefield | Best place to work to start | To much of restriction's | NA | 5 | 5 | 4 | 5 | 3 | |
| 14 | 12 | empReview | 6-Sep-17 | "Software engineer" | 5 | I have been working a | Current Em | Bengaluru | > Good working environm | I dont find any cons | Keep it up | 5 | 5 | 3 | 4 | 3 | |
| 15 | 13 | empReview | 4-Sep-17 | "Good place to work" | 4 | I have been working a | Current Em | Bengaluru | work life balance- a lot of | worst management- a l | NA | 4 | 5 | 3 | 3 | 2 | |
| 16 | 14 | empReview | 17-Sep-17 | "Good company to work f | 3 | I have been working a | Current Em | Bengaluru | Good company to work fo | salary, growth, opportu | please stop say | 4 | 4 | 3 | 2 | 3 | |
| 17 | 15 | empReview | 7-Oct-17 | "No stability" | 1 | I worked at Tesco Ben | Former Em | Bengaluru | Good campusGood payNe | The company constantl | The redundanc | 3 | 1 | 1 | 3 | 1 | |
| 18 | 16 | empReview | 13-Sep-17 | "Technical Program Mana | 3 | I have been working a | Current Em | Bengaluru | A good place to workGood | Nothing as suchProvide | If the role has b | 4 | 3 | 4 | 4 | 2 | |
| 19 | 17 | empReview | 28-Aug-17 | "Best Place to work!! Hap | 5 | I have been working a | Current Em | NA | So many pros to list down | Only one thing i would | Awesome :) | 4 | 5 | 4 | 3 | 5 | |
| 20 | 18 | empReview | 27-Aug-17 | "NOC Engineer worked for | 5 | I worked at Tesco Ben | Former Em | Bengaluru | Amazing company. Life an | Terminates employees c | NA | 5 | 5 | 5 | 4 | 4 | |
| 21 | 19 | empReview | 26-Aug-17 | "Overall Good" | 4 | I have been working a | Current Em | Bengaluru | Work Life balance, Flexibl | Promotions are now a g | Avoid Gold pla | 4 | 4 | 3 | 4 | 3 | |
| 22 | 20 | empReview | 20-Aug-17 | "Graduate software engin | 4 | I worked at Tesco Ben | Former Em | NA | good work life balance in t | not good for freshers.nc | NA | NA | NA | NA | NA | NA | |
| 23 | 21 | empReview | 30-Aug-17 | "Don't join if you want sta | 3 | I have been working a | Current Em | Bengaluru | Good work life balance as | Employees are expecte | Just give a glan | 4 | 3 | 2 | 3 | 1 | |
| 24 | 22 | empReview | 31-Aug-17 | "Overall a good environm | 3 | I have been working a | Current Em | Bengaluru | Well maintained campus, | Managers are neither tr | Kindly promot | 4 | 3 | 2 | 2 | 1 | |
| 25 | 23 | empReview | 29-Aug-17 | "Business Ana;yst" | 3 | I worked at Tesco Ben | Former Em | NA | Great good exposure to bu | No cons no cons no cor | No advice.... ju | NA | NA | NA | NA | NA | |
| 26 | 24 | empReview | 29-Aug-17 | "Property" | 3 | I worked at Tesco Ben | Former Em | Bengaluru | Good campus and facilitie | Transportation from cit | NA | 3 | 4 | 3 | 3 | 2 | |
| 27 | 25 | empReview | 18-Sep-17 | "Manager" | 1 | I have been working a | Current Em | NA | As a company with its facil | Mangers will be just co | You should list | 1 | 3 | 2 | 1 | 1 | |
| 28 | 26 | empReview | 9-Aug-17 | "Good To Work" | 4 | I have been working a | Current Em | Bengaluru | Work and personal Life ba | Need to work on regula | NA | NA | NA | NA | NA | NA | |
| 29 | 27 | empReview | 22-Aug-17 | "TESCO HSC Review" | 2 | I have been working a | Current Em | Bangalore f | Good for freshers, new tec | work life balancemanag | nothing, mana | 1 | 2 | 3 | 2 | 2 | |
| 30 | 28 | empReview | 6-Aug-17 | "Good place to start your | 4 | I worked at Tesco Ben | Former Em | NA | You get to work on (and le | Work Timings, Commute (this doesn't a | | 3 | 4 | 4 | 4 | 3 | |
| 31 | 29 | empReview | 7-Sep-17 | "Pathetic Company. Focu | 1 | I have been working a | Current Em | Bengaluru | Good infrastructure.Bette | Hire and fireIn last 2 ye | What to advice | 3 | 1 | 1 | 3 | 1 | |
| 32 | 30 | empReview | 4-Aug-17 | "Professional place to wo | 4 | I have been working a | Current Em | NA | Very professionalEmployee | Limited growth of the c | Good company | 3 | 5 | 4 | 4 | 4 | |
| 33 | 31 | empReview | 1-Aug-17 | "Senior Officer - Financial | 4 | I worked at Tesco Ben | Former Em | Bengaluru | TESCO is one of the best-es | Employee benefits such | NA | 4 | 5 | 4 | 3 | 4 | |
| 34 | 32 | empReview | 14-Aug-17 | "Tesco test test review" | 3 | I worked at Tesco Ben | Former Em | Bengaluru | Good company. Good and | Bad culture, bad work, NA | NA | 4 | 3 | 3 | 4 | 3 | |

Website Data in Rows and Columns - Salaries

| | | | | | | | |
|-------------------------------|---------|--------------|-------------|------------|------------|-----------|-----------|
| Tesco | Samples | Salary_Range | Weights | Min_Salary | Max_Salary | Min_Units | Max_Units |
| Software Engineer | 69 | 540,002 | 37,260,138 | 375 | 760 | k | k |
| Senior Software Engineer | 55 | 830,937 | 45,701,535 | 472 | 1.4 | k | M |
| Principal Software Engineer | 38 | 1,123,789 | 42,703,982 | 120 | 1.73 | k | M |
| Graduate Software Engineer | 13 | 471,329 | 6,127,277 | 365 | 725 | k | k |
| Assistant Manager | 13 | 1,072,298 | 13,939,874 | 840 | 1.33 | k | M |
| Senior Analyst | 13 | 733,631 | 9,537,203 | 390 | 1.23 | k | M |
| Associate Software Engineer | 12 | 503,279 | 6,039,348 | 130 | 900 | k | k |
| Business Analyst | 12 | 981,949 | 11,783,388 | 335 | 1.95 | k | M |
| Software Engineer - Monthly | 11 | 333,996 | 3,673,956 | 14 | 45 | k | k |
| Officer | 11 | 305,964 | 3,365,604 | 206 | 542 | k | k |
| Team Lead | 10 | 906,336 | 9,063,360 | 500 | 1.77 | k | M |
| | | | | | | | |
| HSBC | Samples | Salary_Range | Weights | Min_Salary | Max_Salary | Min_Units | Max_Units |
| Software Engineer | 156 | 559,599 | 87,297,444 | 240 | 1.2 | k | M |
| Senior Software Engineer | 124 | 904,748 | 112,188,752 | 250 | 1.89 | k | M |
| Analyst | 88 | 815,431 | 71,757,928 | 240 | 1.45 | k | M |
| Business Analyst | 62 | 944,923 | 58,585,226 | 337 | 1.65 | k | M |
| Assistant Manager | 45 | 922,502 | 41,512,590 | 450 | 1.55 | k | M |
| Consultant Specialist | 35 | 1,337,726 | 46,820,410 | 892 | 1.75 | k | M |
| Assistant Manager - Analytics | 35 | 1,014,375 | 35,503,125 | 600 | 1.38 | k | M |
| Manager | 33 | 1,452,161 | 47,921,313 | 692 | 2.45 | k | M |
| Assistant Manager Operations | 31 | 666,300 | 20,655,300 | 441 | 969 | k | k |
| Vice President | 24 | 2,901,019 | 69,624,456 | 1.93 | 3.93 | M | M |
| Team Leader | 24 | 1,378,917 | 33,094,008 | 472 | 1.95 | k | M |
| Software Engineer - Monthly | 23 | 400,320 | 9,207,360 | 17 | 49 | k | k |
| Associate | 22 | 584,224 | 12,852,928 | 250 | 1.8 | k | M |
| Project Manager | 21 | 2,166,936 | 45,505,656 | 1.12 | 3 | M | M |

Tesco's overall compensation is lower than its immediate competitors, also its rating is the least amongst all



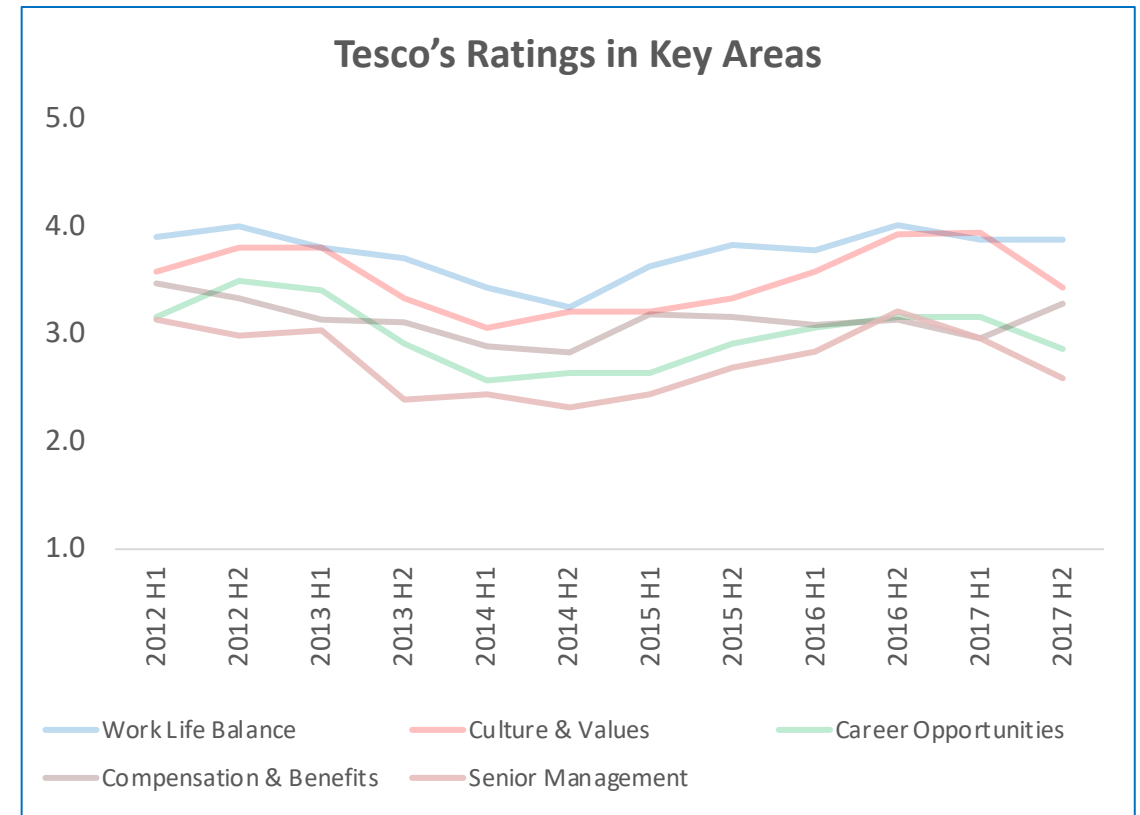
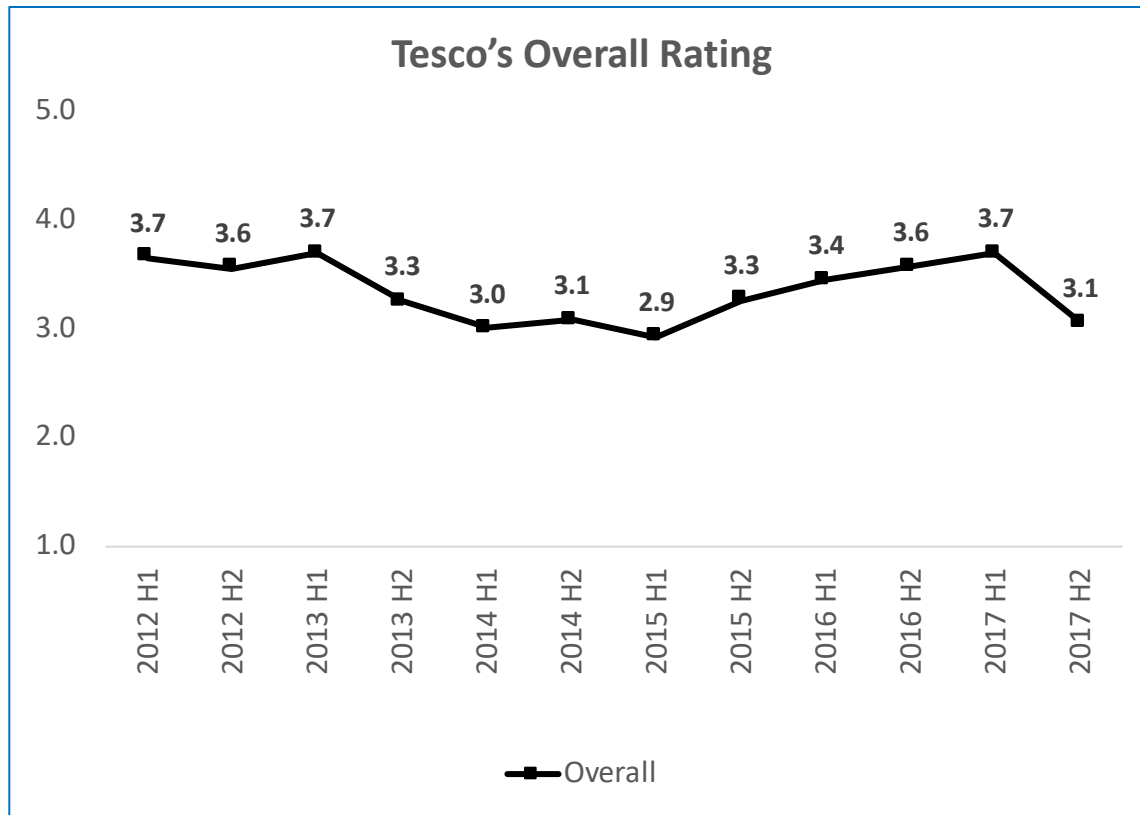
**Tesco Rating
3.3**

**HSBC Rating
3.5**

**Target Rating
3.7**

**Amazon Rating
3.7**

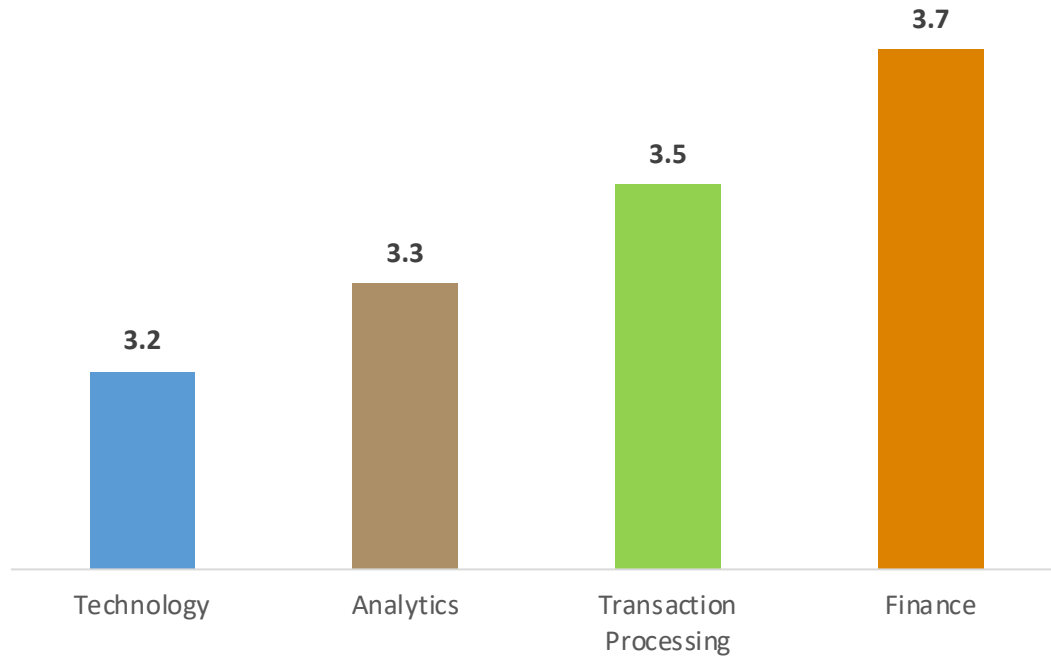
Overall rating has continuously improved since 2015, however it declined drastically in the past few months



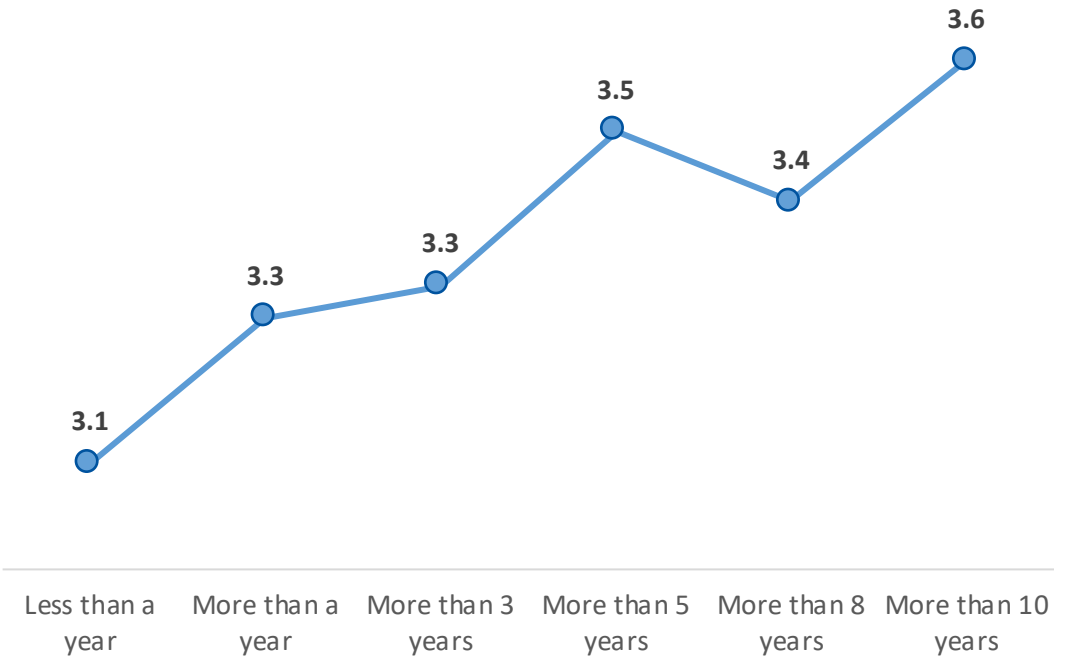
- Career opportunities and senior management have a higher impact on the overall rating in comparison with compensation/benefits and work life balance

Employee ratings improve with work experience; there is scope for improvement in Technology and Analytics job functions

Employee Ratings vs Function



Employee Ratings vs Work Experience



Planning to Cluster reviews to get the big picture ?

Forming TF-IDF Matrix

```
from sklearn.feature_extraction.text import TfidfVectorizer

vectorizer = TfidfVectorizer()

# CSR Matrix below will have the tf-idf vectors for all the Pros
csr_mat = vectorizer.fit_transform(pros_cleaned_list)
print(vectorizer.get_feature_names())
print(csr_mat.shape)
print(type(csr_mat))

# Create a DataFrame aligning labels and titles: df
df_clusters = pd.DataFrame({'cluster': labels, 'article_clean': pros_cleaned_list})

# Display df sorted by cluster label
print(df_clusters.sort_values('cluster'))
```

Dimensional Reduction followed by Clustering

```
# Dimensional reduction removes noise. Try with different count vectorizer parameters
# Clustering helps to identify topics. Try with different count vectorizer parameters

from sklearn.decomposition import TruncatedSVD
from sklearn.cluster import KMeans
from sklearn.pipeline import make_pipeline

# Create a TruncatedSVD instance: svd
svd = TruncatedSVD(n_components = 100)

# Create a KMeans instance: kmeans
kmeans = KMeans(n_clusters=8)

# Create a pipeline: pipeline
pipeline = make_pipeline(svd,kmeans)

# Fit the pipeline to articles
pipeline.fit(csr_mat)

# Calculate the cluster labels: labels
labels = pipeline.predict(csr_mat)

# Create a DataFrame aligning labels and titles: df
df_clusters = pd.DataFrame({'cluster': labels, 'article_clean': pros_cleaned_list})

# Display df sorted by cluster label
print(df_clusters.sort_values('cluster'))
```

Sample set of Pros

| |
|---|
| Great good exposure to business. |
| Good campus and facilities to employees |
| As a company with its facilities for employees is good but managers do not let employee |
| Work and personal Life balance is 100% |
| Good for freshers, new technologies to work on |
| You get to work on (and learn) the latest technologies, Facilities/Amenities, Gender ratio, Lots of activities (like college), My managers were awesome |
| Good infrastructure.Better pay.Flexibility in work depending on team. |
| Very professionalEmployee friendly |

Simple count works well here...

Methodology

1. Identify the key words that could define a topic
2. Map reviews to topic if there is at least one key word in the review
3. One to many mapping is thus possible

For example, adjacent image is collection of all positive reviews on “Management”. But they are also the positive reviews on “Work Culture” and “Job Satisfaction”

Pros

> Good working environment> Supportive management> Excellent Transport facility

1. Good employee benefits.2. Great management.

Management good; HR good; Floor space excellent

Transparent management, good culture, good managers, excellent learning

Very good culture. People focused & encourages work life balance-

Process oriented. Emphasis on structured thinking and standardisation-

Friendly working environment Good management Excellent Work-Life balance

Making rapid transformation into product management space . Hence IT is currently a happening place both in terms of technology

With the change in management , so many good opportunities are provided for employee growth and employees are given ample

U get the independence to work without management involving to check on the same

Good work atmosphere Supportive management Good work culture

Very professional atmosphere

good breath and opportunity ambitious strong in program management

Work life balanceGood management Innovative opportunities On the whole a good place to work for Flexible for moms Decent hikes

Campus, Top Management is helpful, Growth

Tesco used to be a good place to work with opportunities to specialise in the retail domain and its landmark CDM program. The top

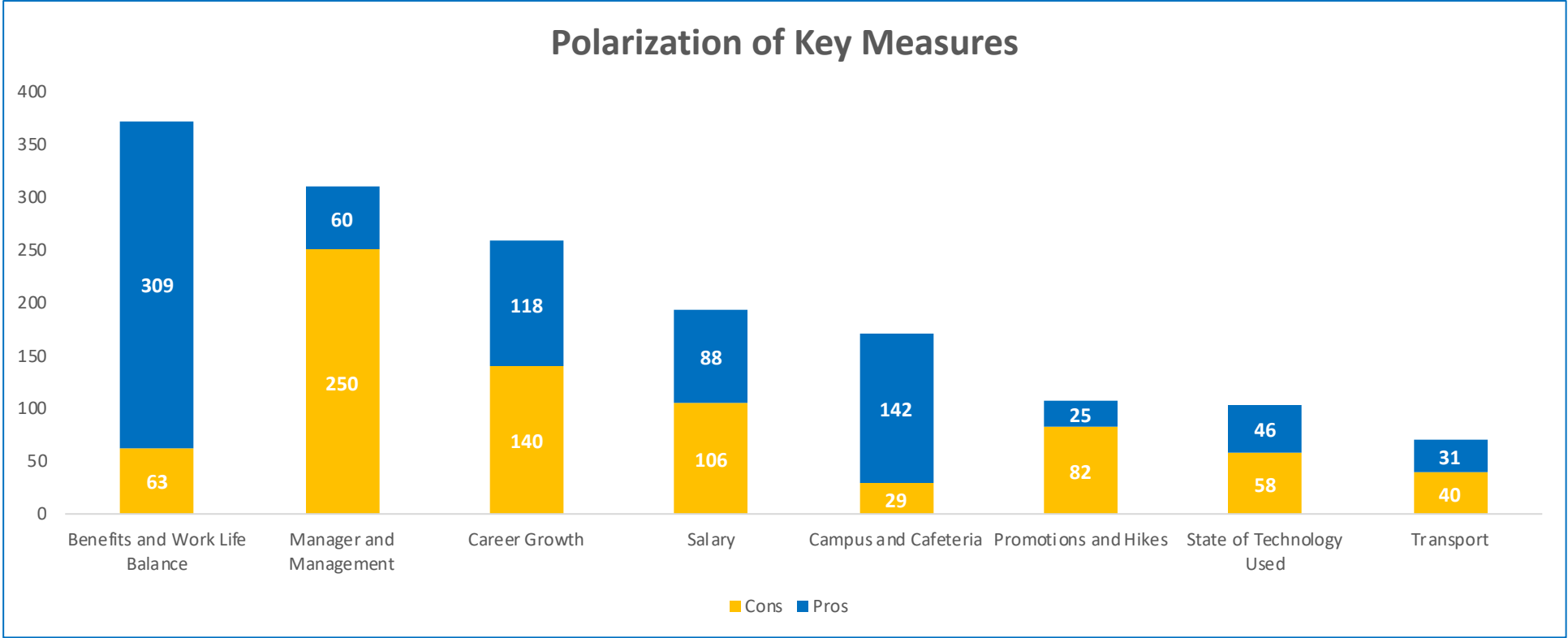
Helpful and Understanding Indian Management teamGood Campus FacilitiesWork Life Balance

You can learn technology; meet few great technical people and learn from them. If you are good at stakeholder management you can thrive.

work life balance, supportive management, work & culture

work culture, learning, management, growth

Management, promotions and hike are the biggest concerns for colleagues; while they are more satisfied with campus facilities, cafeteria and work life balance



Recommendations from topic labels

- We would need to continue the work from home, flexible timings policies
- Improve **communications about** severance package during inevitable circumstances
- Work on communications about **rewards and recognition**
- Building a **grievance tool** for colleagues to escalate work place ethics violations
- Introduce more **transparent structure** on promotions
- **Create opportunities** for all teams to get more exposure in technology

Q&A 😊