

Project 10 : Compensation Analytics (v 1

Objective:

Analyze synthetic HR Compensation dataset to extract key insights: pay distribution, gender pay and bonus allocation.

Business Context:

Compensation & Benefits (C&B) analytics helps HR leaders answer:

- Are we paying fairly across levels, genders, departments?
- How are bonuses distributed?
- Where are the gaps or outliers that need intervention?

This notebook covers:

- 1 . Load & explore dataset
- 2 . Core metrics (Avg/Median CTC)
- 3 . Bonus % analysis
- 4 . Gender pay gap
- 5 . Visuals (CTC by level, gender gap, bonus distribution)
- 6 . Export artifacts

Cloning into 'hr-tech-portfolio'...

remote: Enumerating objects: 1184, done.

remote: Counting objects: 100% (120/120), done.

remote: Compressing objects: 100% (98/98), done.

remote: Total 1184 (delta 75), reused 33 (delta 22), pack-reused 1064 (from 1
Receiving objects: 100% (1184/1184), 11.40 MiB | 28.13 MiB/s, done.

Resolving deltas: 100% (717/717), done.

/content/hr-tech-portfolio/hr-tech-portfolio

✓ Synthetic dataset created at data/employee_compensation_sample.csv

Out[33]:

	EmployeeID	Gender	JobLevel	Department	CTC	Bonus
0	1	Male	3	Finance	1 5 4 7 0 2 1.0	2 8 2 0 5 3
1	2	Female	3	Tech	1 3 5 8 2 9 0.8	1 4 8 0 7 7
2	3	Female	2	Ops	6 6 6 5 7 4.2	1 1 6 1 0 4
3	4	Male	4	Sales	2 0 9 5 6 2 7.0	3 5 2 3 9 4
4	5	Male	3	Ops	1 6 1 2 3 6 3.0	2 3 1 6 3 8

Data Shape: (500, 6)

Out[34]:

	EmployeeID	Gender	JobLevel	Department	CTC							Bonus					
0	1	Male	3	Finance	1	5	4	7	0	2	1.0	2	8	2	0	5	3
1	2	Female	3	Tech	1	3	5	8	2	9	0.8	1	4	8	0	7	7
2	3	Female	2	Ops	6	6	6	5	7	4.2		1	1	6	1	0	4
3	4	Male	4	Sales	2	0	9	5	6	2	7.0	3	5	2	3	9	4
4	5	Male	3	Ops	1	6	1	2	3	6	3.0	2	3	1	6	3	8
5	6	Male	1	Sales	2	7	1	5	5	6.0		3	3	7	1	9	
6	7	Male	5	HR	3	0	4	0	3	0	5.0	5	9	8	7	3	5
7	8	Female	4	Sales	1	8	2	4	9	2	0.1	1	3	0	9	2	8
8	9	Female	5	Finance	2	7	5	0	9	1	8.4	4	3	6	7	7	2
9	10	Female	3	Tech	1	4	1	9	8	6	0.7	1	1	1	7	1	2

Average CTC: ₹ 12.18 Lakhs
Median CTC: ₹ 8.88 Lakhs

Out[35]:

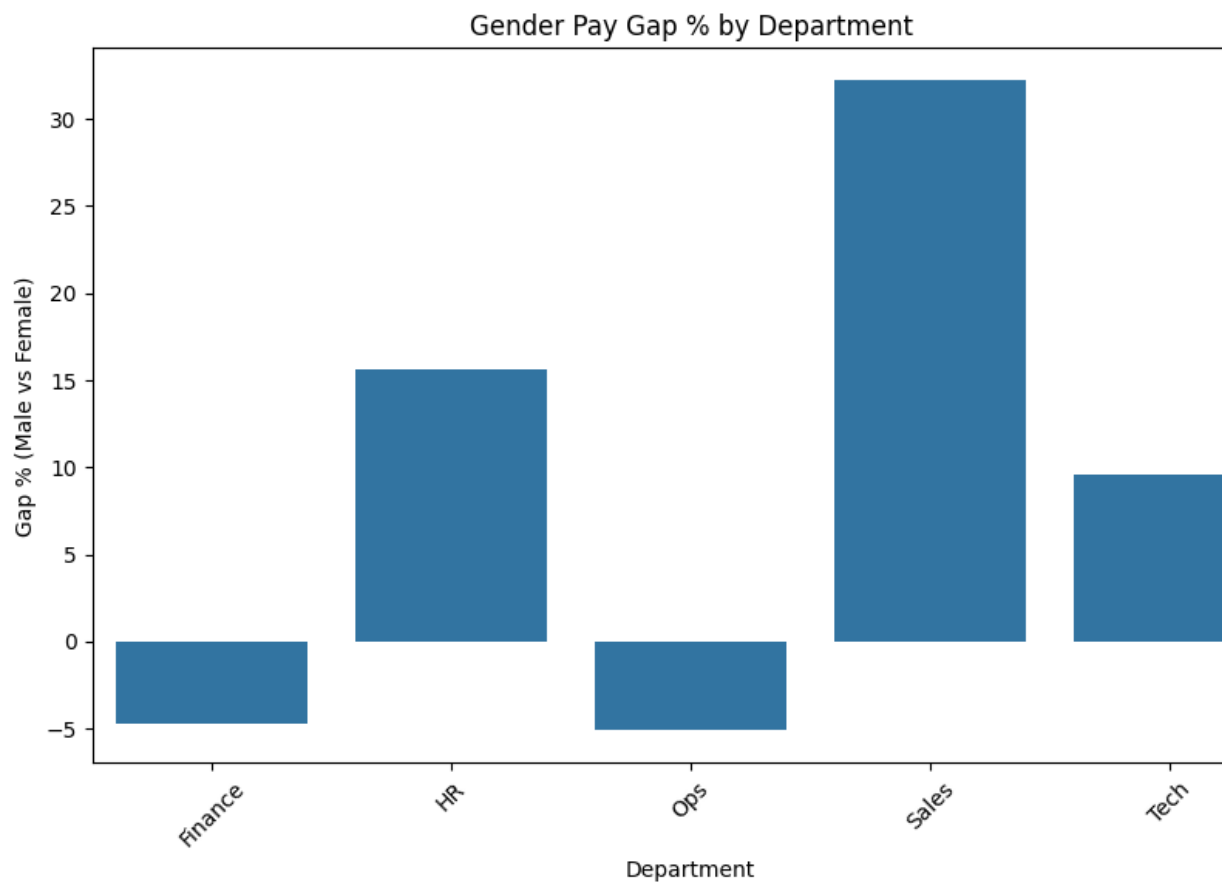
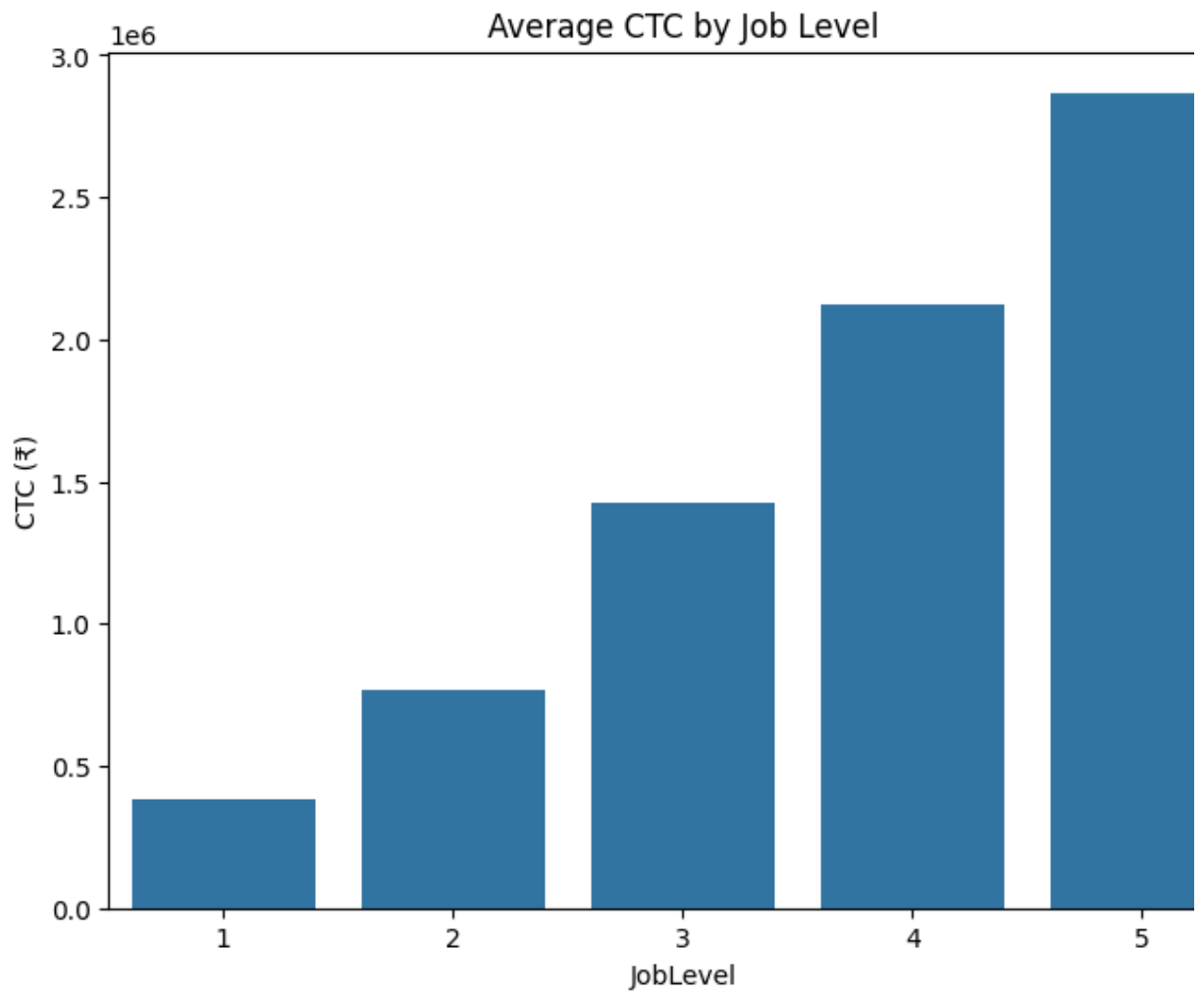
	Department	mean							median						
0	Finance	1	1	0	2	3	0	3	e+06	8	3	3	4	5	6.00
1	HR	1	3	1	5	3	1	5	e+06	9	9	7	4	4	4.00
2	Ops	1	1	3	7	1	8	0	e+06	8	3	0	9	5	7.50
3	Sales	1	1	7	6	3	1	1	e+06	8	4	4	1	6	9.50
4	Tech	1	3	6	0	3	4	5	e+06	1	3	6	0	2	3.45

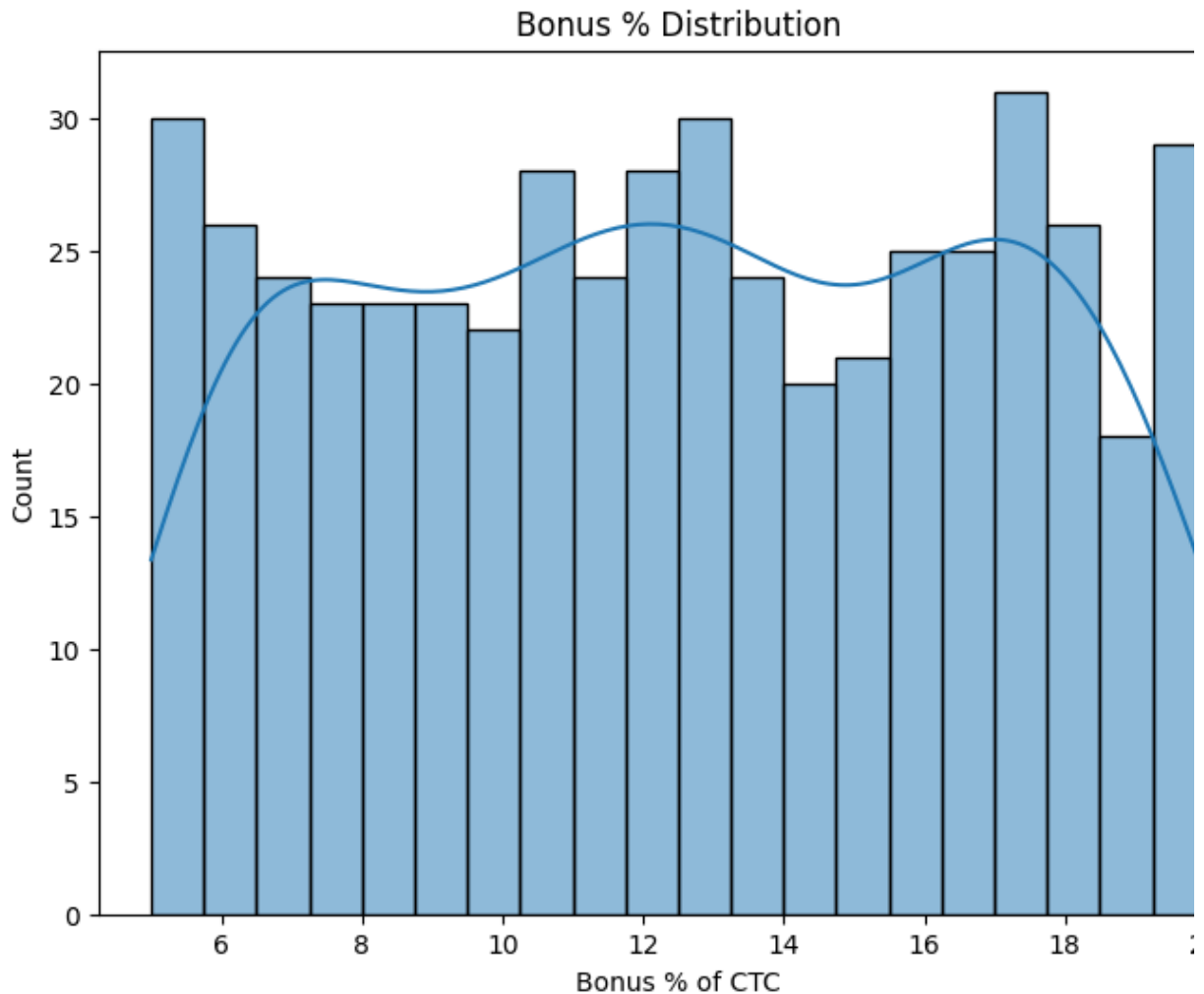
Out[36]:

	EmployeeID					CTC					Bonus					Bonus %
	4 3 4	4 3 5	3 8 0 4 6 4.0				7 6 0 7 6	1 9.9 9 5 5 8 4								
	1 0 6	1 0 7	7 4 2 6 5 0.0				1 4 8 2 0 4	1 9.9 5 6 1 0 3								
	4 8 3	4 8 4	3 1 0 1 2 8.3				6 1 8 4 5	1 9.9 4 1 7 4 7								
	4 0 9	4 1 0	1 0 1 0 0 0 6.0				2 0 1 2 2 2	1 9.9 2 2 8 5 2								
	1 5 3	1 5 4	1 4 4 4 7 2 9.0				2 8 7 1 5 6	1 9.8 7 6 1 1 5								

Out[37]:

Gender	Department	Female										Male													
0	Finance	1	1	2	8	9	5	5	e+0	6	1	0	7	8	2	1	4	e+0	6	-4	7	0	6	0	
1	HR	1	1	7	7	9	2	6	e+0	6	1	3	9	6	4	9	9	e+0	6	1	5	6	5	1	5
2	Ops	1	1	7	6	0	8	7	e+0	6	1	1	1	9	0	9	6	e+0	6	-5	0	9	2	5	
3	Sales	9	3	9	2	3	0	8	e+0	5	1	3	8	5	5	0	0	e+0	6	3	2	2	0	9	9
4	Tech	1	2	8	5	2	6	2	e+0	6	1	4	2	1	2	6	1	e+0	6	9	5	6	8	8	





✅ Artifacts generated:

- data/Comp_Analytics_Processed.csv
- images/comp_ctc_by_joblevel.png
- images/comp_gender_gap.png
- images/comp_bonus_dist.png

Enter your GitHub token:

Enumerating objects: 12, done.

Counting objects: 100% (12/12), done.

Delta compression using up to 2 threads

Compressing objects: 100% (9/9), done.

Writing objects: 100% (9/9), 72.00 KiB | 12.00 MiB/s, done.

Total 9 (delta 3), reused 0 (delta 0), pack-reused 0

remote: Resolving deltas: 100% (3/3), completed with 3 local objects.

To https://github.com/AMBOT-pixel96/hr-tech-portfolio.git

e15ff22..7bbf8f8 main -> main

✅ **Conclusions**

- **Average CTC:** X Lakhs, **Median CTC:** Y Lakhs
- **By Level:** Pay rises steadily with seniority.
- **By Gender:** Gap of Z% in Dept A (highest disparity).
- **Bonuses:** Top 10 % employees capture ~__% of total bonus pool.

📄 Notebook: [Compensation_Analytics_V 1 .ipynb](#)