



oooo

---

# LAYOFF TRENDS ANALYSIS: INSIGHTS & PATTERNS

oooo

# INTRODUCTION

My name is Mansur Ahamad. I have done MCA from Asian International University and Bsc(Hons)Mathematics from Aligarh Muslim University. currently i am doing internship from Unified Mentor. i am working as Data Analyst. i am well in Power Bi, Excel, SQL, Python.

-In this project i utilized sql query to clean data and perform analysis



# TABLE OF CONTENTS

- Schema
- Tables
- Company History
- Clean data
- Layoffs Analysis
- Key Factor



# CLEAN DATA

## Checking Duplicate Values

```
WITH CTE_DUPLICATES AS (
    SELECT *,
        ROW_NUMBER() OVER (PARTITION BY company, location, industry, total_laid_off, stage,
                            ORDER BY date) AS ROW_NO
    FROM layoffs_staging
)
SELECT *
FROM CTE_DUPLICATES
WHERE ROW_NO > 1;
```

	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country	funds_raised_millions	ROW_NO
1	Airlift	Lahore	Logistics	NULL	1	7/12/2022	Series B	Pakistan	109	2
2	Airy Rooms	Jakarta	Travel	NULL	1	5/7/2020	Unknown	Indonesia	NULL	2
3	Arrival	London	Transportation	NULL	0.3	7/13/2022	Post-IPO	United Kingdom	629	2
4	Aura Financial	SF Bay Area	Finance	NULL	NULL	4/14/2020	Unknown	United States	584	2
5	Automox	Boulder	Infrastructure	NULL	NULL	6/13/2022	Series C	United States	152	2
6	Better.com	New York City	Real Estate	NULL	NULL	8/26/2022	Unknown	United States	905	2
7	Bybit	Singapore	Crypto	NULL	NULL	6/20/2022	Unknown	Singapore	NULL	2
8	Casper	New York City	Retail	NULL	NULL	9/14/2021	Post-IPO	United States	339	2

Result 24 X

# CLEAN DATA

# Remove Duplicate Values

```
CREATE TABLE `world_layoffs`.`layoffs_staging2`(`company` text,`location` text,`industry` text,  
`total_laid_off` INT,`percentage_laid_off` text,`date` text,`stage` text,`country` text,  
`funds_raised_millions` int,ROW_NO INT  
);  
select * from layoffs_staging2  
insert into layoffs_staging2  
select *, row_number() over(partition by company,location,industry,total_laid_off,stage,country)  
AS ROW_NO from layoffs_staging;  
select * from layoffs_staging2 where ROW_NO > 1  
delete from layoffs_staging2 where ROW_NO > 1
```

<a href="#">Result Grid</a>		<a href="#">Filter Rows:</a>	<input type="text"/>	<a href="#">Export:</a>		<a href="#">Wrap Cell Content:</a>	
-----------------------------	--	------------------------------	----------------------	-------------------------	--	------------------------------------	--

# STANDARDIZED DATA

```
SELECT DISTINCT company, TRIM(company) FROM layoffs_staging2;  
UPDATE layoffs_staging2  
SET company = TRIM(company);
```

Result Grid | Filter Rows:

	company	TRIM(company)
▶	Included Health	Included Health
	&Open	&Open
	#Paid	#Paid
	100 Thieves	100 Thieves
	10X Genomics	10X Genomics

# STANDARDIZED DATA

```
SELECT industry FROM layoffs_staging2 WHERE industry LIKE 'Crypto%';
UPDATE layoffs_staging2
SET industry = 'Crypto' |
WHERE industry LIKE 'Crypto%';
```

industry
Crypto

# STANDARDIZED DATA

```
SELECT DISTINCT country FROM layoffs_staging2 ORDER BY country;
```

country
United Arab Emirates
United Kingdom
United States
United States.

```
SELECT DISTINCT country, TRIM(TRAILING '.' FROM country) AS cleaned_country  
FROM layoffs_staging2  
ORDER BY country;
```

```
UPDATE layoffs_staging2  
SET country = TRIM(TRAILING '.' FROM country);
```

country
United Arab Emirates
United Kingdom
United States

# STANDARDIZED DATA

```
SELECT DISTINCT date FROM layoffs_staging2 ;  
SELECT date, STR_TO_DATE(date, '%m/%d/%Y') AS converted_date  
FROM layoffs_staging2 ;  
UPDATE layoffs_staging2  
SET date = STR_TO_DATE(date, '%m/%d/%Y');  
ALTER TABLE layoffs_staging2 MODIFY COLUMN date DATE;
```

Result Grid |

	date
▶	2022-07-25
	2022-11-17
	2023-01-27
	2022-07-13
	2022-08-04

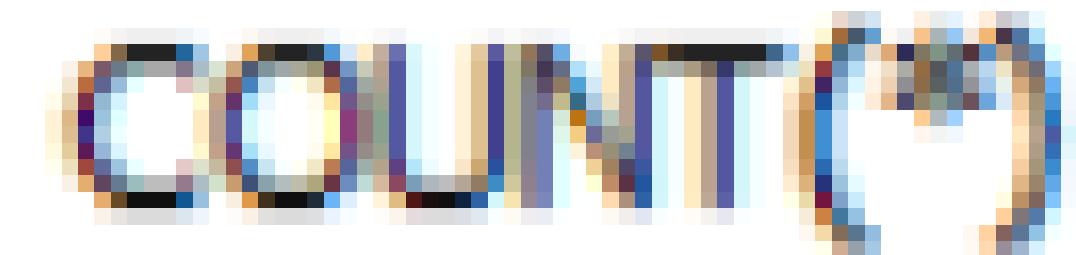
# HANDLE NULL VALUES

```
SELECT * FROM layoffs_staging2 WHERE total_laid_off IS NULL  
AND percentage_laid_off IS NULL;
```

Result Grid | Filter Rows:

	total_laid_off	percentage_laid_off
▶	NULL	NULL
	NULL	NULL

```
SELECT COUNT(*) FROM layoffs_staging2  
WHERE total_laid_off IS NULL  
AND percentage_laid_off IS NULL;
```



338

# HANDLE NULL VALUES

```
DELETE FROM layoffs_staging2  
WHERE total_laid_off IS NULL  
AND percentage_laid_off IS NULL;
```

Result Grid	
	COUNT(*)
→	0

# HANDLE NULL VALUES

```
SELECT COUNT(*) FROM layoffs_staging2  
WHERE industry IS NULL  
OR industry = '';
```

Result Grid	
	COUNT(*)
↓	4

# HANDLE NULL VALUES

```
SELECT table1.company, table1.industry AS missing_industry,  
       table2.industry AS correct_industry  
  FROM layoffs_staging2 AS table1  
 JOIN layoffs_staging2 AS table2  
    ON table1.company = table2.company  
 WHERE (table1.industry IS NULL OR table1.industry = '')  
   AND table2.industry IS NOT NULL AND table2.industry <> '';
```

Result Grid | Filter Rows:

	company	missing_industry	correct_industry
▶	Airbnb		Travel
	Airbnb		Travel
	Carvana		Transportation
	Carvana		Transportation
	Juul		Consumer

# HANDLE NULL VALUES

```
UPDATE layoffs_staging2
JOIN layoffs_staging2 AS table2
ON layoffs_staging2.company = table2.company
SET layoffs_staging2.industry = table2.industry
WHERE layoffs_staging2.industry IS NULL OR layoffs_staging2.industry = ''
AND table2.industry IS NOT NULL AND table2.industry <> '';
```

```
SELECT COUNT(*) FROM layoffs_staging2
WHERE industry IS NULL
OR industry = '';
```

Result Grid

COUNT(*)
0

# HANDLE NULL VALUES

```
SELECT ROW_NO from layoffs_staging2
```

| Result Grid | 

	ROW_NO
▶	NULL
	NULL

```
ALTER TABLE layoffs_staging2 DROP COLUMN ROW_NO;  
select ROW_NO from layoffs_staging2
```

76 16:50:27 SELECT ROW\_NO from layoffs\_staging2 LIMIT 0, 2000

Error Code: 1054. Unknown column 'ROW\_NO' in 'field list'

# ANALYSIS

## TOP 3 COMPANY LAYOFFS

```
SELECT company, industry, SUM(total_laid_off) AS Highest_layoff  
FROM layoffs_staging2  
GROUP BY company, industry  
ORDER BY Highest_layoffs DESC  
LIMIT 3;
```

Result Grid | Filter Rows:  Export: Wrap Cell Content:

	company	industry	Highest_layoffs
▶	Amazon	Retail	36300
	Google	Consumer	24000
	Meta	Consumer	22000

# ANALYSIS

## TOP 3 COMPANY highest percentage layoffs

```
SELECT company,
       ROUND(SUM(CAST(REPLACE(percentage_laid_off, '%', '') AS FLOAT)), 2) AS laid_off_Per
  FROM layoffs_staging2
 GROUP BY company
 ORDER BY laid_off_Per DESC
 LIMIT 3;
```

Result Grid | Filter Row

	company	laid_off_Per
▶	Pollen	3.28
▶	Rubius	3.14
▶	Zeus Living	2.52

# ANALYSIS

## TOP 3 COMPANY layoff percentage per year

```
SELECT
    company,
    YEAR(date) AS layoff_year,
    ROUND(SUM(CAST(REPLACE(percentage_laid_off, '%', '') AS FLOAT)), 2) AS total_layoff_percentage
FROM layoffs_staging2
GROUP BY company, layoff_year
ORDER BY total_layoff_percentage DESC
LIMIT 3;
```

Result Grid | Filter Rows:

	company	layoff_year	total_layoff_percentage
▶	Rubius	2022	3.14
	Pollen	2022	2.66
	BlockFi	2022	2.4

# ANALYSIS

## TOP 3 COMPANY cumulative layoffs Yr/month

SELECT

```
DATE_FORMAT(date, '%Y-%m') AS layoff_month,  
SUM(total_laid_off) AS monthly_layoffs,  
SUM(SUM(total_laid_off)) OVER (ORDER BY DATE_FORMAT(date, '%Y-%m')) AS cumulative_layoffs  
FROM layoffs_staging2  
GROUP BY layoff_month  
ORDER BY layoff_month;
```

Result Grid | Filter Rows:  | Export

	layoff_month	monthly_layoffs	cumulative_layoffs
	2020-05	51608	125190
	2020-06	15254	140444
	2020-07	14224	154668
	2020-08	2938	157606
	2020-09	1218	158824

# ANALYSIS

## % contribution of company in total layoffs,

SELECT

```
company,  
SUM(total_laid_off) AS company_layoffs,  
ROUND(SUM(total_laid_off) * 100.0 / (SELECT SUM(total_laid_off) FROM layoffs_staging2), 2) AS percentage_contribution  
FROM layoffs_staging2  
GROUP BY company  
ORDER BY percentage_contribution DESC;
```

Result Grid | Filter Rows:  Export:

	company	company_layoffs	percentage_contribution
▶	Amazon	36300	4.76
	Google	24000	3.15
	Meta	22000	2.89
	Microsoft	20000	2.62
	Philips	20000	2.62



# THANK YOU

## **CONTACT DETAILS**



Noida



7217206241



[mansurahamad945@gmail.com](mailto:mansurahamad945@gmail.com)

