

- Null values in total_laid_off column

```
146 • select * from layoffs_staging2
147 where total_laid_off is null
```

1 / 12

Result Grid Filter Rows: <input type="text"/> Export: Wrap Cell Content:					
	company	location	industry	total_laid_off	percentage_laid_off
▶	E Inc.	Toronto	Transportation	NULL	NULL
	Included Health	SF Bay Area	Healthcare	NULL	0.06
	100 Thieves	Los Angeles	Retail	NULL	NULL
	2U	Washington D.C.	Education	NULL	0.2
	5B Solar	Sydney	Energy	NULL	0.25
	80 Acres Farms	Cincinnati	Food	NULL	0.1
	98point6	Seattle	Healthcare	NULL	0.1
	Accolade	Seattle	Healthcare	NULL	NULL
	ActiveCampaign	Chicago	Marketing	NULL	0.15
	Ada	Toronto	Support	NULL	NULL
	Adara	SF Bay Area	Travel	NULL	NULL

- In Industry column as well, there are null values

```
151 • select * from layoffs_staging2
152 where industry is null
153 or industry='';
154
```

Result Grid Filter Rows: <input type="text"/> Export: Wrap Cell Content:								
	company	location	industry	total_laid_off	percentage_laid_off	date	stage	country
▶	Airbnb	SF Bay Area		30	NULL	2023-03-03	Post-IPO	United States
	Bally's Interactive	Providence	NULL	NULL	0.15	2023-01-18	Post-IPO	United States
	Carvana	Phoenix		2500	0.12	2022-05-10	Post-IPO	United States
	Juul	SF Bay Area		400	0.3	2022-11-10	Unknown	United States

- Taking 'Carvana' company as an example, industry is Transportation, hence we can populate the blank row with the same value.

```
157 • select * from layoffs_staging2
158     where company = 'Carvana';
159
```

company	location	industry	total_laid_off	percentage_laid_off	date	stage	country
Carvana	Phoenix		2500	0.12	2022-05-10	Post-IPO	United States
Carvana	Phoenix	Transportation	NULL	NULL	2023-01-13	Post-IPO	United States
Carvana	Phoenix	Transportation	1500	0.08	2022-11-18	Post-IPO	United States

- We can use self-join where conditions are company and location should be the same.

```
162 • select * from layoffs_staging2 t1
163     join layoffs_staging2 t2
164     on t1.company=t2.company
165     and t1.location=t2.location
166     where (t1.industry is null or t1.industry='')
167     and t2.industry is not null
168
```

company	location	industry	total_laid_off	percentage_laid_off	date	stage	country
Airbnb	SF Bay Area		30	NULL	2023-03-03	Post-IPO	United States
Airbnb	SF Bay Area		30	NULL	2023-03-03	Post-IPO	United States
Carvana	Phoenix		2500	0.12	2022-05-10	Post-IPO	United States
Carvana	Phoenix		2500	0.12	2022-05-10	Post-IPO	United States
Carvana	Phoenix		2500	0.12	2022-05-10	Post-IPO	United States
Juul	SF Bay Area		400	0.3	2022-11-10	Unknown	United States
Juul	SF Bay Area		400	0.3	2022-11-10	Unknown	United States

- Update all blank values to NULL

```
2 • update layoffs_staging2
3     set industry=null
4     where industry='';
5
6     -- check
7 • select * from layoffs_staging2
8     where industry is null
9     or industry='';
```

company	location	industry	total_laid_off	percentage_laid_off	date	stage	country
Airbnb	SF Bay Area	NULL	30	NULL	2023-03-03	Post-IPO	United States
Bally's Interactive	Providence	NULL	NULL	0.15	2023-01-18	Post-IPO	United States
Carvana	Phoenix	NULL	2500	0.12	2022-05-10	Post-IPO	United States
Juul	SF Bay Area	NULL	400	0.3	2022-11-10	Unknown	United States

- We need populate t2.industry values in t1.industry

```

173 • select t1.industry, t2.industry from layoffs_staging2 t1
174      join layoffs_staging2 t2
175      on t1.company=t2.company
176      and t1.location=t2.location
177      where (t1.industry is null or t1.industry='')
178      and t2.industry is not null;

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	industry	industry		
▶	NULL	Travel		
	NULL	Transportation		
	NULL	Transportation		
	NULL	Consumer		

- Updating t1 and joining t2 where company is the same. Setting t1.industry to t2.industry values. t1.industry is blank or null now and t2.industry is not null

update layoffs_staging2 t1

join layoffs_staging2 t2

on t1.company=t2.company

set t1.industry=t2.industry

where t1.industry is null

and t2.industry is not null;