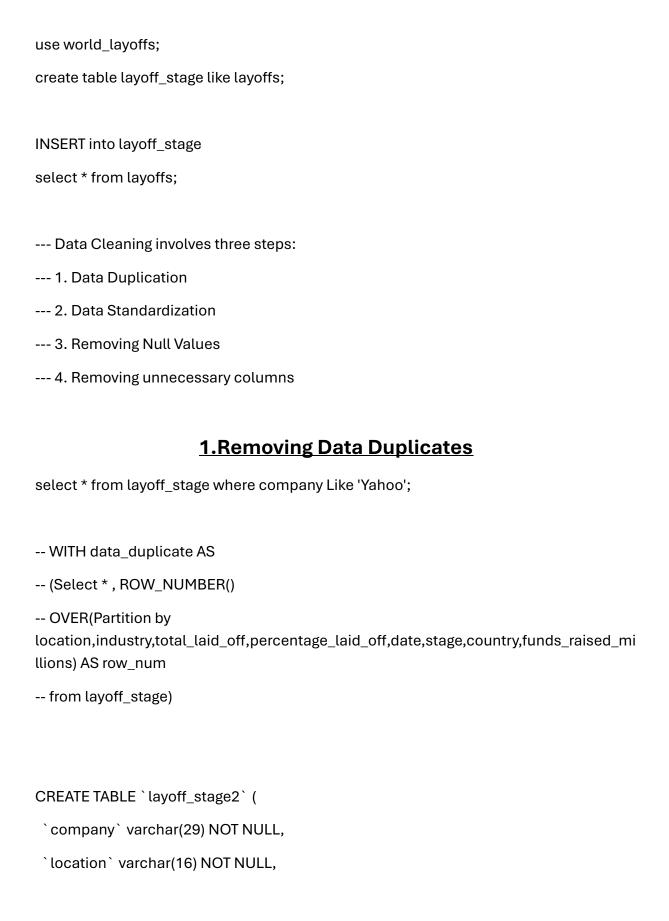
Data Cleaning SQL File



```
`industry` varchar(15) DEFAULT NULL,
```

) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci;

Insert into layoff_stage2

Select *, ROW_NUMBER()

OVER(Partition by

location,industry,total_laid_off,percentage_laid_off,date,stage,country,funds_raised_mi llions) AS row_num

from layoff_stage;

Select * from layoff_stage2 where row_num>1;

delete from layoff_stage2 where row_num>1;

2. Standardizing Data

UPDATE layoff_stage2 SET company = TRIM(company);

SELECT Industry from layoff_stage2 where industry like 'Crypto%';

UPDATE layoff_stage2 SET INDUSTRY = "Crypto" where industry like 'Crypto%';

[`]total_laid_off` int DEFAULT NULL,

[`]percentage_laid_off` decimal(6,4) DEFAULT NULL,

[`]date` date DEFAULT NULL,

[`]stage` varchar(14) DEFAULT NULL,

[`]country` varchar(20) NOT NULL,

[`]funds_raised_millions` decimal(10,4) DEFAULT NULL,

[`]row_num` INT

```
SELECT DISTINCT(Industry) from layoff_stage2;
```

SELECT DISTINCT(Location) from layoff_stage2;

SELECT DISTINCT(Country) from layoff_stage2 order by Country;

UPDATE layoff_stage2 SET Country = "United States" where Country like 'United S%';

SELECT (Date) from layoff_stage2;

3. Removing Null Values

Select * from layoff_stage2 where company is null;

Select * from layoff_stage2 where industry is null;

Select * from layoff_stage2 where total_laid_off is null and percentage_laid_off is null;

Select * from layoff_stage2 where stage is null or stage = 'unknown';

Select * from layoff_stage2 where funds_raised_millions is null;

Select * from layoff_stage2 as l1

Join layoff_stage2 as l2

on l2.company=l1.company and l1.location=l2.location

where (l1.industry is NULL or l1.industry = "") and l2.industry is not null;

Update layoff stage2 as l1

Join layoff_stage2 as l2

on l2.company=l1.company and l1.location=l2.location

SET l1.industry = l2.industry

where (l1.industry is NULL or l1.industry = "") and l2.industry is not null;

4. Deleting unnecessary rows

Delete from layoff_stage2 where total_laid_off is null and percentage_laid_off is null;

ALTER Table layoff_stage2 drop column row_num;