

# Sales Analysis of Store by mysql



Srinivas golla

12 min read · Just now

[Open in app](#)

Search



-srinivas Golla



In today's fast-paced world of technology, having organized databases and skilled teams to understand them is crucial for efficiency. While getting things done is important, having high-quality information is even more critical for better results. With the market becoming more competitive, it's essential to analyze our products, regions, categories, and customer segments. This analysis will help us find patterns and insights, enabling us to optimize efficiency and increase profits. Remember, providing value to our customers is the key to success.

We'll conduct our analysis using Excel and SQL. Excel will store our data initially, and SQL will help us make sense of that data. The sales performance analysis will follow these Data Analysis steps: Ask questions, Prepare the data, Process the information, Analyze the results, Share insights, and Act upon the findings.

We'll clarify the business challenge presented to us, which can be understood as "Identifying the optimal products, regions, categories, and customer segments for the Superstore to focus on or steer clear of, with the aim of boosting profitability."

### **Business Goals:**

- Boost Profits: Explore strategies to increase our overall profitability.
- Detect Trends: Pinpoint emerging patterns within our data.
- Transform Insights into Advice: Utilize identified insights to formulate practical recommendations.

### **Deliverables:**

- Concise Business Objectives Summary
- Thorough Documentation: Covering all data cleaning, manipulation, and analysis steps.
- Recommendations: Actionable advice derived from our analysis and insights.

### **Prepare:**

- . The data is publicly available through Kaggle under  
<https://www.kaggle.com/datasets/vivek468/superstore-dataset-final>
- . It comes with 9995 rows with 9994 being pure data and the other one row being the column headers. It contains data recorded between the 3rd of January 2014 (the first order date) to the 5th of January 2018 (the last shipping date). (The last order date is the 30th of December 2017, so we will instead use the order dates range to represent our 4 years of business).
- . It contains the data of 793 customers.
- . The data contains the 21 columns namely; Row ID, Order ID, Order Date, Ship Date, Ship Mode, Customer ID , Customer Name, Segment, Postal Code, City, State,

Country, Region, Product ID, Category, Sub-Category, Product Name, Sales, Quantity, Discount and Profit.

### 1. Data Overview:

- Open the CSV file in Excel to get an initial look at our data.
- Check for any missing data using conditional formatting.

Remove Duplicates:

- . Eliminate any duplicate rows from the dataset to maintain accuracy.

Column Formatting:

- Ensure all columns are correctly formatted for seamless SQL analysis.

While exploring the dataset, we made a few observations:

- Data Consistency:
  - The data appears correct and consistent, requiring minor edits for enhanced analysis.
- Conditional Formatting:
  - Used conditional formatting to highlight blank cells in yellow, confirming the absence of missing values.
- Duplicate Check:
  - Employed the ‘remove duplicates’ command, revealing no identical rows with all parameters for any customer.
- Date and Currency Formatting:
  - Verified proper formatting of Order Date and Ship Date as dates.
  - Formatted Sales and Profits columns as currencies, considering their financial nature.
  - Discount column remains a percentage, to be formatted later by multiplying it by 100.

## Analyze:

For the analysis part, we will string out the most important components of our data to answer our business objectives.

Let's load our data into SQL and check the first 5 rows to make sure it imported well.

```
select * from store limit 5;
```

Row_ID	Order_ID	Order_Date	Ship_Date	Ship_Mode	Customer_ID	Customer_Name	Segment	Country	City	State	Postal_Code	Region
1	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South
2	CA-2016-152156	2016-11-08	2016-11-11	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South
3	CA-2016-138688	2016-06-12	2016-06-16	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	California	90036	West
4	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South
5	US-2015-108966	2015-10-11	2015-10-18	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South

Okay, let's perform an exploratory data analysis with our input on the superstore dataset. A list of tasks will be answered followed by the query input and query result. At the end of our analysis, we will transition to a dashboard reflecting the answers to the most important components that will solve the business problem posed upon us.

1. What are total sales and total profits of each year ?

```
select round(sum(sales),4) as sales ,year(order_date) as year from store group
```

	sales	year
▶	601265.261	2016
	464426.236	2015
	481763.7981	2014
	724994.5612	2017

- In 2014, the total sales amounted to \$481,763.80.
- The year 2015 saw a slight decrease in sales to \$464,426.24.

- Sales experienced a notable increase in 2016, reaching \$601,265.26.
- The highest sales were recorded in 2017, totaling \$724,994.56.

```
select round(sum(profit),4) as profit ,year(order_date) as year from store group by year
```

	profit	year
▶	80130.6524	2016
	60907.6897	2015
	49044.4263	2014
	92774.9858	2017

- In 2014, the profit was \$49,044.43.
- In 2015, the profit increased to \$60,907.69.
- The year 2016 saw a further rise in profit to \$80,130.65.
- The highest profit was recorded in 2017, totaling \$92,774.99.

## 2. What are the total profits and total sales per quarter?

```
select round(sum(sales),2)as sales,round(sum(profit),2)as profit ,quarter(order_date) as quarter from store group by quarter
```

	sales	profit	quater	year
▶	232650.39	37286.21	4	2016
	134672.65	16097.08	2	2016
	180559.67	23376.67	4	2015
	86283.22	11238.18	2	2014
	132358.09	15685.76	2	2017
	179072.02	21546.7	4	2014
	142601.19	12641.58	3	2014
	193621.16	26251.77	3	2017
	128609.75	16505.86	3	2015
	91971.55	11316.25	1	2016
	276153.48	27291.11	4	2017
	141970.66	15431.11	3	2016
	86626.14	11825.31	2	2015
	68630.68	9199.86	1	2015
	122861.83	23546.35	1	2017
	73807.36	3617.96	1	2014

The dataset captures sales and profit details for different quarters over several years. Each entry includes sales and profit figures, the specific quarter (1 to 4), and the corresponding year. This information provides a comprehensive view of performance metrics, enabling a nuanced analysis of quarterly trends and annual variations.

### 3. What region generates the highest sales and profits ?

```
select round(sum(sales),4) as sales,round(sum(profit),4)as profit,region from
```

	sales	profit	region
▶	388983.585	46035.689	South
	497800.8728	40128.903	Central
	672194.054	90672.0127	East
	713471.3445	106021.1495	West

it's evident that the West region plays a crucial role in both sales and profits. With the highest sales figure of \$713,471.34 and the highest profit of \$106,021.15 among all regions, the West region stands out as a key contributor to overall business success. It would be advisable for the client to focus on and potentially invest further in strategies that capitalize on the strengths and opportunities in the West region, maximizing its potential for sustained growth and profitability.

#### 4. What state and city brings in the highest sales and profits ?

```
select round(sum(sales),2)as sales,round(sum(profit),2)as profit ,city,state f
```

	sales	profit	city	state
▶	255248.97	61624.06	New York City	New York
	173168.87	29806.92	Los Angeles	California
	117772.58	28869	Seattle	Washington
	110917.04	17176.67	San Francisco	California
	107486.47	-13732.34	Philadelphia	Pennsylvania

The crux suggests that New York City leads in both sales and profits, while Philadelphia shows a negative profit, indicating a potential area for improvement or optimization. Analyzing the sales and profit data at the city and state level can guide strategic decisions for enhancing overall performance.

#### 5. The relationship between discount and sales and the total discount per category?

```
select round(avg(sales),2) as average_sales,discount from store group by discou
select round(sum(discount),2)as total_discount ,category from store group by ca
```

	average_sales	discount
▶	892.71	0.5
	578.4	0.1
	567.54	0.4
	536.79	0.32
	529.97	0.15
	498.63	0.45
	454.74	0.3
	230.36	0
	213.22	0.2
	102.36	0.7
	58.63	0.8
	47.25	0.6

- The data suggests that higher discounts (0.5, 0.1, 0.4) generally result in higher average sales.
- Consider experimenting with moderate discounts (around 0.2 to 0.4) to optimize sales without compromising too much on revenue.
- Extreme discounts (0.7, 0.8) have a significant impact on reducing average sales. Be cautious and strategic when implementing deep discounts.
- Evaluate the relationship between discount and profitability to ensure that higher sales do not come at the cost of overall profit margins.

A balanced approach to discounts, focusing on moderate ranges, can be a strategic pricing plan to maximize both sales and profitability. Regularly analyze and adjust the strategy based on customer response and market dynamics.

	total_discount	category
▶	899.3	Office Supplies
	361.69	Furniture
	243.2	Technology

- The highest total discount is in the “Office Supplies” category, indicating a potentially more competitive pricing strategy or promotional focus in this category.
- “Furniture” has the second-highest total discount, suggesting there might be room to explore pricing strategies or promotions to boost sales in this category.
- “Technology” has the lowest total discount, which could imply that the pricing strategy for technology products may be more conservative or that these products are perceived as more premium.

## 6. What category generates the highest sales and profits in each region and state ?

```
select round(sum(sales),2) as total_sales,category,region,state from store group by category,region,state
select round(sum(profit),2) as total_profit,category,region,state from store group by category,region,state
```

The image shows two side-by-side MySQL query results. Both queries are titled 'Result Grid' and have 'Filter Rows:' and 'Export' buttons.

**Left Query (Sales):**

	total_sales	category	region	state
▶	159183.17	Technology	West	California
▶	153836.07	Furniture	West	California
	137548.35	Office Supplies	West	California
	127453.53	Technology	East	New York
	92661.57	Furniture	East	New York

**Right Query (Profit):**

	total_profit	category	region	state
▶	42172.7	Technology	East	New York
	36733.4	Office Supplies	West	California
	29440.37	Technology	West	California
	25761.83	Office Supplies	East	New York
	15016.65	Technology	West	Washington

top 5 sales and profit by category

- Regional Sales Comparison: California in the West region outperforms New York in the East region across all three product categories.
- Category Performance: Within California, Technology products lead in sales, followed by Furniture and Office Supplies.
- East vs. West Dynamics: The West region demonstrates higher combined sales, possibly influenced by regional demand patterns or market preferences.
- Insights for Strategy: Consider tailoring marketing or product offerings based on regional preferences. Explore opportunities for boosting sales in specific categories or regions. Regularly analyze and adapt strategies to align with evolving market dynamics.

7. What subcategory generates the highest sales and profits in each region and state ?

```
select round(sum(sales),2) as total_sales,Sub_Category,region,state from store
select round(sum(profit),2) as total_profit,Sub_Category,region,state from stor
```

Result Grid					Filter Rows:	E	Result Grid					Filter Rows:	Export:	
	total_sales	Sub_Category	region	state		total_profit	Sub_Category	region	state		total_profit	Sub_Category	region	state
▶	67876.62	Phones	West	California		17320.08	Machines	East	New York		17320.08	Machines	East	New York
	63079.37	Chairs	West	California		13385.1	Phones	East	New York		13385.1	Phones	East	New York
	47472.65	Phones	East	New York		11095.66	Accessories	West	California		11095.66	Accessories	West	California
	46634.25	Chairs	East	New York		11070.55	Binders	Central	Michigan		11070.55	Binders	Central	Michigan
	45430.23	Tables	West	California		11028.35	Binders	East	New York		11028.35	Binders	East	New York
	43183.5	Machines	East	New York		9734.06	Binders	West	California		9734.06	Binders	West	California
	42635.13	Storage	West	California		9442.42	Copiers	West	Washington		9442.42	Copiers	West	Washington
	37255.01	Accessories	West	California		8849.97	Copiers	Central	Indiana		8849.97	Copiers	Central	Indiana
	32261.42	Binders	East	New York		7889.85	Copiers	West	California		7889.85	Copiers	West	California
	29492.02	Machines	West	California		7553.36	Paper	West	California		7553.36	Paper	West	California
	28589.57	Phones	Central	Texas		7170.47	Storage	West	California		7170.47	Storage	West	California
	27718.72	Binders	West	California		7111.01	Chairs	East	New York		7111.01	Chairs	East	New York
	27558.52	Bookcases	West	California		6992.23	Appliances	West	California		6992.23	Appliances	West	California
	26810.39	Storage	East	New York		6279.45	Phones	West	California		6279.45	Phones	West	California
	26572.45	Chairs	Central	Texas		6041.28	Binders	Central	Minnesota		6041.28	Binders	Central	Minnesota
	25062.32	Tables	West	Washington		5966.6	Binders	West	Washington		5966.6	Binders	West	Washington
	24573.94	Machines	South	Florida		5839.95	Copiers	East	New York		5839.95	Copiers	East	New York
	24559.52	Copiers	West	California		5717.31	Binders	South	Georgia		5717.31	Binders	South	Georgia

- . Focus on Profitable Sub-Categories: Given the higher profits, consider emphasizing or expanding offerings in sub-categories like Machines and Phones in the East region.
  - . Regional Profit Analysis: Analyze profit patterns in different regions. For instance, assess if Accessories are performing well in the West region and adjust marketing strategies accordingly.
  - . Explore Opportunities in Binders: While Binders show consistent profitability, explore opportunities to optimize sales and marketing strategies in regions like Michigan and New York.
  - . Customer Demand Insights: Investigate customer preferences and demand in each region to align product offerings with local market trends.
  - . Regular Monitoring: Continuously monitor and adjust strategies based on changing market dynamics and customer preferences to maximize overall profitability.
8. What are the names of the products that are the most and least profitable to us?

```
select product_name ,profit from store where profit =(select max(profit) from s
select product_name ,profit from store where profit =(select min(profit) from s
```

Result Grid		Filter Rows:	Export:
	product_name	profit	
▶	Canon imageCLASS 2200 Advanced Copier	8399.976	

Result Grid		Filter Rows:	Export:	Wrap
	product_name	profit		
▶	Cubify CubeX 3D Printer Double Head Print	-6599.978		

- The Cubify CubeX 3D Printer Double Head Print is showing a substantial loss, amounting to -\$6,599.978.
- On the other hand, the Canon imageCLASS 2200 Advanced Copier is the most profitable, with a profit of \$8,399.976.
- Assess the pricing strategy and production costs for the Cubify CubeX printer. Consider whether adjustments can be made to improve profitability or if it aligns with broader business goals.
- Given the high profitability of the Canon imageCLASS 2200 Advanced Copier, consider strategic marketing efforts to promote and upsell similar high-margin products.

## 9. What segment makes the most of our profits and sales ?

```
select segment ,sum(sales) as total_sales from store group by segment order by
select segment ,sum(profit) as total_profit from store group by segment order b
```

Result Grid		Filter Rows:
	segment	total_sales
▶	Consumer	1150166.1819999903
	Corporate	696604.5138000002
	Home Office	425679.1605000003

Result Grid		Filter Rows:
	segment	total_profit
▶	Consumer	132669.77519999977
	Corporate	90366.3009000002
	Home Office	59821.678100000085

- Consumer Focus: Continue emphasizing products that appeal to the Consumer segment, as it is a major driver of both sales and profits.
- Corporate Profitability: Explore strategies to improve profitability within the Corporate segment, such as tailored promotions or cost-effective solutions.
- Home Office Growth: Identify growth opportunities within the Home Office segment to increase both sales and profitability.
- Segment-Specific Marketing: Tailor marketing efforts to each segment's preferences and needs, ensuring a targeted and effective approach.

10. How many customers do we have (unique customer IDs) in total and how much per region and state?

```
select count(distinct customer_id) as total_customers from store;
select count(distinct customer_id) as total_customers ,region,state from store
```

	total_customers	region	state
▶	573	West	California
	413	East	New York
	369	Central	Texas
	248	East	Pennsylvania
	234	Central	Illinois
	218	West	Washington
	198	East	Ohio
	180	South	Florida
	121	South	North Carolina
	106	South	Virginia
	104	Central	Michigan
	99	West	Arizona
	83	South	Tennessee
	83	South	Georgia
	74	West	Colorado
	69	Central	Indiana
	62	East	Massachusetts
	61	East	New Jersey
	58	South	Kentucky
	51	Central	Wisconsin

- West (California, Washington, Arizona, Oregon, Nevada, New Mexico):

- High customer count suggests strong market presence.
- Suggestion: Capitalize on the existing customer base by offering tailored promotions or expanding product offerings.
- **East (New York, Pennsylvania, Ohio, Massachusetts, New Jersey):**
  - Moderate to high customer count, indicating a significant market share.
  - Suggestion: Strengthen customer engagement through targeted marketing and explore opportunities for sales growth.
- **Central (Texas, Illinois, Michigan, Minnesota, Nebraska):**
  - Balanced customer count across states.
  - Suggestion: Optimize regional strategies to enhance sales and potentially tap into new markets within the Central region.
- **South (Florida, North Carolina, Virginia, Georgia, Tennessee):**
  - Moderate customer count, representing a substantial market.
  - Suggestion: Implement strategies to increase market share, such as localized promotions or enhanced customer service.

## Key Suggestions to Improve Superstore Performance:

1. **Regional Targeting:** Tailor marketing and sales strategies based on the characteristics and preferences of each region. Consider region-specific promotions to increase customer engagement.
2. **Customer Retention:** Implement customer loyalty programs to retain existing customers. Focus on personalized communication and exclusive offers to enhance customer loyalty.
3. **Market Expansion:** Explore opportunities to enter new markets within existing regions or consider expanding into regions with lower customer counts. Conduct market research to understand local demands.
4. **Product Diversification:** Analyze the product mix in each region and introduce or emphasize products based on regional preferences. This can help boost sales by aligning offerings with customer needs.

5. E-commerce Enhancement: Strengthen the online presence and e-commerce capabilities, especially in regions with lower customer counts. Invest in user-friendly online platforms and seamless delivery services.
6. Customer Feedback: Gather customer feedback to understand satisfaction levels and identify areas for improvement. Use this information to refine services, address customer concerns, and enhance overall customer experience.
11. Average shipping time per class and in total?

```
select Ship_Mode,avg(ship_date-order_date) as average_time ,count(ship_mode) as
```



The screenshot shows the MySQL Workbench interface with a query editor containing the following SQL code:

```
select Ship_Mode,avg(ship_date-order_date) as average_time ,count(ship_mode) as
```

Below the code, the results are displayed in a grid:

	Ship_Mode	average_time	count
▶	Same Day	0.0455	527
	First Class	59.3897	1501
	Second Class	116.9873	1886
	Standard Class	195.0304	5780

1. Efficiency Improvement: Implement measures to enhance overall logistics efficiency, aiming for quicker delivery times across all classes.
2. Promotion of Same Day Delivery: If feasible, promote and incentivize Same Day Delivery to increase its adoption, especially for time-sensitive orders.
3. Customer Communication: Clearly communicate delivery expectations to customers, allowing them to make informed choices. Transparency builds trust and satisfaction.
4. Technology Integration: Explore technology solutions, such as route optimization software, to streamline delivery processes and reduce transit times.
5. Customer Preferences Analysis: Analyze customer preferences and expectations regarding delivery times. Tailor logistic strategies to align with customer needs and market trends.

6. Cost-Benefit Analysis: Evaluate the cost implications of faster delivery options against potential revenue gains and customer satisfaction. Balance the trade-offs to find an optimal logistic strategy.

## summary

### 1. Sales by Region:

- *Crux:* West region dominates in both sales and profits.
- *Action:* Focus on strengthening offerings and promotions in the West region. Evaluate and optimize strategies in other regions.

### 2. Sales by Category:

- *Crux:* Technology products lead in sales; Furniture shows potential for improvement.
- *Action:* Enhance marketing efforts for Furniture. Regularly update product offerings based on market trends.

### 3. Profit by Segment:

- *Crux:* Consumer segment contributes significantly to both sales and profits.
- *Action:* Continue consumer-focused strategies. Optimize profitability in Corporate and Home Office segments.

### 4. Discount and Average Sales:

- *Crux:* Higher discounts generally lead to higher average sales.
- *Action:* Strategically implement moderate discounts. Regularly assess the impact on overall profitability.

### 5. Sub-Category Profit:

- *Crux:* Machines and Phones show high profits.
- *Action:* Emphasize marketing and promotions for high-profit sub-categories.

### 6. Product Profitability:

- *Crux:* Cubify CubeX faces significant losses; Canon imageCLASS is highly profitable.
- *Action:* Reevaluate the pricing and strategy for Cubify CubeX. Promote high-profit products like Canon imageCLASS.

## 7. Customer Count by Region:

- *Crux:* West region has the highest customer count.
- *Action:* Strengthen existing customer base in the West. Explore market expansion in regions with lower customer counts.

## 8. Logistic Delivery Analysis:

- *Crux:* Varying delivery times across shipping modes.
- *Action:* Optimize logistic processes for efficiency. Promote faster delivery options and enhance overall communication.

## General Recommendations:

- Regularly monitor market trends and customer feedback.
- Invest in technology for data analysis and process optimization.
- Conduct periodic cost-benefit analyses for strategic decision-making.
- Implement customer-centric initiatives for improved satisfaction and loyalty.

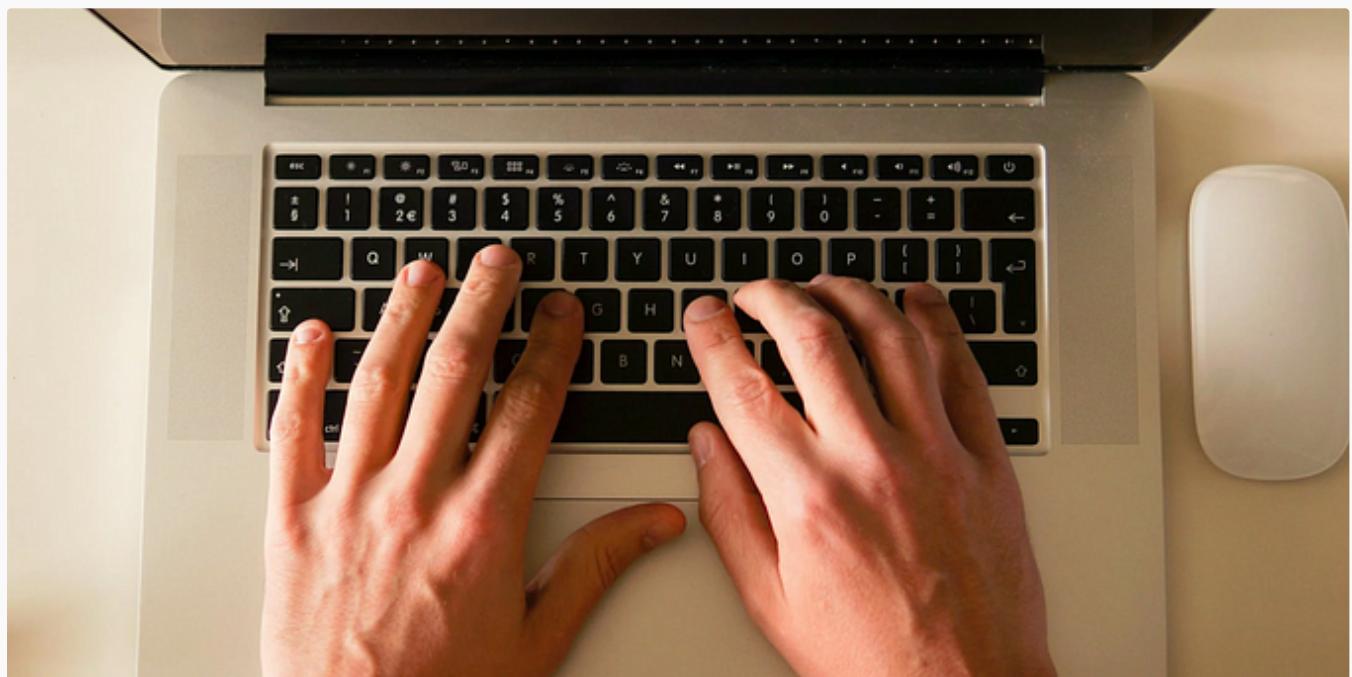
These actions aim to strategically position the Superstore for sustained growth, profitability, and enhanced customer satisfaction in the dynamic market landscape. Regularly reassess and adapt strategies based on evolving market dynamics and customer preferences.

[Edit profile](#)

## Written by Srinivas golla

0 Followers

## Recommended from Medium



 SQL Fundamentals in Stackademic

### E-commerce Customer Churn Analysis Using SQL

Decoding Customer Behavior with SQL

◆ · 2 min read · 6 days ago

 72

 1



...



 Alex Deng in The Airbnb Tech Blog

## Airbnb at KDD 2023

KDD (Knowledge and Data Mining) is a flagship conference in data science research. Hosted annually by a special interest group of the...

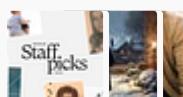
10 min read · 2 days ago

 46  2



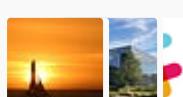
...

## Lists



### Staff Picks

541 stories · 564 saves



### Stories to Help You Level-Up at Work

19 stories · 378 saves



### Self-Improvement 101

20 stories · 1087 saves



### Productivity 101

20 stories · 992 saves

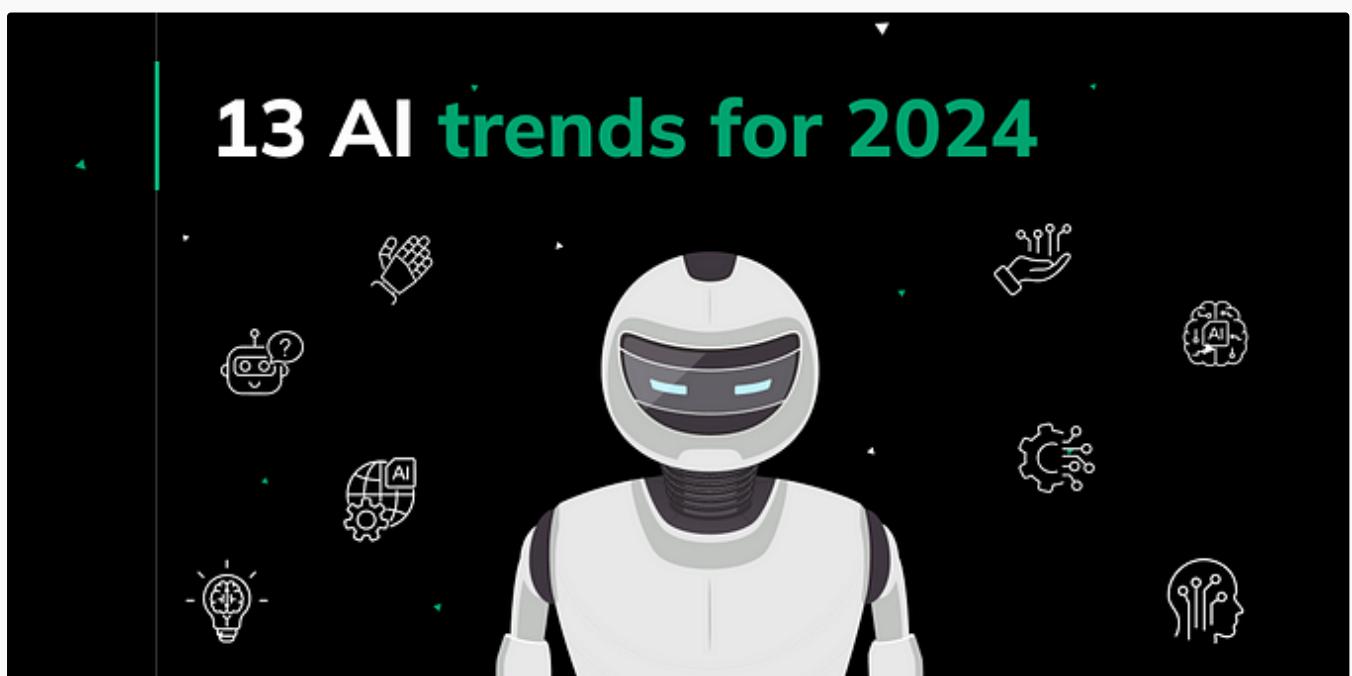


 Syed Hamed Raza in Dev Genius

## Unveiling the Power of MySQL Date and Time Functions

A Practical Guide

7 min read · 2 days ago



 Pragmatic Coders

## AI predictions: Top 13 AI trends for 2024

Explore the future with our comprehensive guide to the top 13 AI trends anticipated for 2024.

13 min read · Dec 10

👏 2.4K

💬 62



...



## Hord Launches an X Ambassador Program

 NyamCatCrypto

### **Hord Ambassador Program: Earn up to 500\$/month**

Embark on an exciting crypto journey that promises both enjoyment and rewards through the Hord Ambassador Program. Influencers are invited...

2 min read · 6 days ago

👏 280

💬 2



...

 Paul Rose

## I Found The Best Side Hustle for Food Lovers

How to Make Money Writing About Food

8 min read · 4 days ago

 608 16

...

See more recommendations