Databases II
Semester 2025-III
Workshop No. 1 — Project Definition and Database Modeling
Personalized E-Commerce Platform

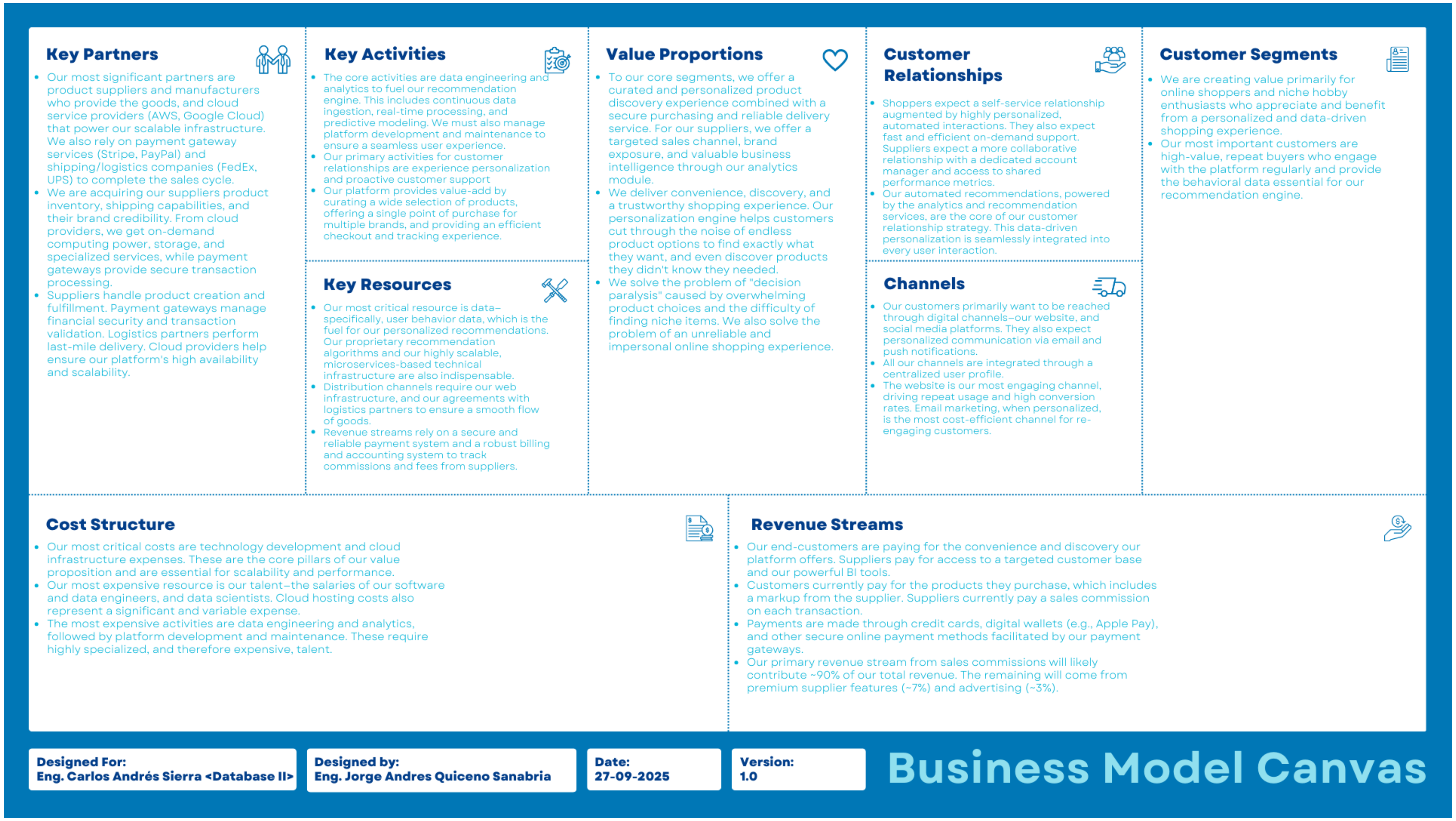
Eng. Jorge Andrés Quiceno S.[1], Eng. Laurent David Chaverra Córdoba[2]
and Eng. Juan David Olmos Corredor[3]

September 28, 2025

Contents

1 Business Model Canvas

The Business Model Canvas describes the value proposition, customer segments, channels, revenue streams, and other key aspects of the application.



Key Partners

- Our most significant partners are product suppliers and manufacturers who provide the goods, and cloud service providers (AWS, Google Cloud) that power our scalable infrastructure. We also rely on payment gateway services (Stripe, PayPal) and shipping/logistics companies (FedEx, UPS) to complete the sales cycle.
- We are acquiring our suppliers product inventory, shipping capabilities, and their brand credibility. From cloud providers, we get on-demand computing power, storage, and specialized services, while payment gateways provide secure transaction processing.
- Suppliers handle product creation and fulfillment. Payment gateways manage financial security and transaction validation. Logistics partners perform last-mile delivery. Cloud providers help ensure our platform's high availability and scalability.

Key Activities

- The core activities are data engineering and analytics to fuel our recommendation engine. This includes continuous data ingestion, real-time processing, and predictive modeling. We must also manage platform development and maintenance to ensure a seamless user experience.
- Our primary activities for customer relationships are experience personalization and proactive customer support
- Our platform provides value-add by curating a wide selection of products, offering a single point of purchase for multiple brands, and providing an efficient checkout and tracking experience.

Key Resources

- Our most critical resource is data–specifically, user behavior data, which is the fuel for our personalized recommendations. Our proprietary recommendation algorithms and our highly scalable, microservices-based technical infrastructure are also indispensable.
- Distribution channels require our web infrastructure, and our agreements with logistics partners to ensure a smooth flow of goods.
- Revenue streams rely on a secure and reliable payment system and a robust billing and accounting system to track commissions and fees from suppliers.

Value Proportions

- To our core segments, we offer a curated and personalized product discovery experience combined with a secure purchasing and reliable delivery service. For our suppliers, we offer a targeted sales channel, brand exposure, and valuable business intelligence through our analytics module.
- We deliver convenience, discovery, and a trustworthy shopping experience. Our personalization engine helps customers cut through the noise of endless product options to find exactly what they want, and even discover products they didn't know they needed.
- We solve the problem of "decision paralysis" caused by overwhelming product choices and the difficulty of finding niche items. We also solve the problem of an unreliable and impersonal online shopping experience.

Customer Relationships

- Shoppers expect a self-service relationship augmented by highly personalized, automated interactions. They also expect fast and efficient on-demand support. Suppliers expect a more collaborative relationship with a dedicated account manager and access to shared performance metrics.
- Our automated recommendations, powered by the analytics and recommendation services, are the core of our customer relationship strategy. This data-driven personalization is seamlessly integrated into every user interaction.

Channels

- Our customers primarily want to be reached through digital channels–our website, and social media platforms. They also expect personalized communication via email and push notifications.
- All our channels are integrated through a centralized user profile.
- The website is our most engaging channel, driving repeat usage and high conversion rates. Email marketing, when personalized, is the most cost-efficient channel for re-engaging customers.

Customer Segments

- We are creating value primarily for online shoppers and niche hobby enthusiasts who appreciate and benefit from a personalized and data-driven shopping experience.
- Our most important customers are high-value, repeat buyers who engage with the platform regularly and provide the behavioral data essential for our recommendation engine.

Cost Structure

- Our most critical costs are technology development and cloud infrastructure expenses. These are the core pillars of our value proposition and are essential for scalability and performance.
- Our most expensive resource is our talent–the salaries of our software and data engineers, and data scientists. Cloud hosting costs also represent a significant and variable expense.
- The most expensive activities are data engineering and analytics, followed by platform development and maintenance. These require highly specialized, and therefore expensive, talent.

Revenue Streams

- Our end-customers are paying for the convenience and discovery our platform offers. Suppliers pay for access to a targeted customer base and our powerful BI tools.
- Customers currently pay for the products they purchase, which includes a markup from the supplier. Suppliers currently pay a sales commission on each transaction.
- Payments are made through credit cards, digital wallets (e.g., Apple Pay), and other secure online payment methods facilitated by our payment gateways.
- Our primary revenue stream from sales commissions will likely contribute ~90% of our total revenue. The remaining will come from premium supplier features (~7%) and advertising (~3%).

Designed For:
Eng. Carlos Andrés Sierra <Database II>

Designed by:
Eng. Jorge Andres Quiceno Sanabria

Date:
27-09-2025

Version:
1.0

Business Model Canvas

Figure 1: Business Model Canvas for the E-Commerce Platform

2 Requirements

2.1 Functional Requirements

Each functional requirement (FR) describes essential user and system behaviors.

2.1.1 FR1 – User & Identity

FR1.1 **User registration & authentication** – allow secure registration, login, and identity verification.
Acceptance: Users can create accounts, verify their identity, and log in securely.

FR1.2 **User profiles & preferences** – manage shipping addresses, payment references, and personalization preferences.
Acceptance: Profile changes are reflected across the system consistently and promptly.

2.1.2 FR2 – Product Catalog & Inventory

FR2.1 **Catalog management** – create, update, delete, and search products, categories, and attributes.
Acceptance: Updates appear in catalog views within defined refresh times.

FR2.2 **Inventory management** – update inventory counts in real-time and prevent overselling.
Acceptance: Inventory levels adjust immediately when purchases occur.

2.1.3 FR3 – Shopping & Orders

FR3.1 **Cart, checkout & order lifecycle** – support shopping cart, and checkout.
Acceptance: Orders are consistently recorded and visible in user history.

FR3.2 **Returns & refunds** – handle return requests and update order history.
Acceptance: Return actions update both user-facing history and internal records.

2.1.4 FR4 – Personalization & Recommendations

FR4.1 **Personalized recommendations** – provide tailored product suggestions based on user activity.
Acceptance: Recommendations are generated within defined latency limits.

FR4.2 **A/B testing** – support experiments for personalization effectiveness.
Acceptance: Experiment results are captured and available for analysis.

2.1.5 FR5 – Data Ingestion & Processing

FR5.1 **Event collection** – capture user interactions, orders, and transactions.
Acceptance: Events are reliably ingested and stored for downstream processing.

FR5.2 **Data processing** – support both real-time and batch processing for analytics and features.
Acceptance: Processed data is consistent with raw events and usable in reporting.

2.1.6 FR6 – Search & Analytics

FR6.1 **Search & filtering** – provide full-text and faceted search capabilities.
Acceptance: Search returns accurate results within target latency thresholds.

FR6.2 **Business intelligence dashboards** – provide managerial insights into sales, inventory, and performance.
Acceptance: Dashboards reflect up-to-date data according to freshness SLAs.

2.1.7 FR7 – Observability & Governance

FR7.1 **System monitoring** – log key events, track metrics, and generate alerts.
Acceptance: System operators receive alerts for critical failures and performance degradations.

FR7.2 **Data governance** – enforce schema consistency and enable data lineage tracking.
Acceptance: Data consumers can verify schema and trace lineage for critical datasets.

2.2 Non-Functional Requirements

These define quality attributes and measurable targets.

2.2.1 Performance & Latency

- Event ingestion: typical events are processed end-to-end within a few seconds.
- Recommendation responses: generated within 50–200 ms under normal load.
- Search queries: typical results within 200–500 ms.

2.2.2 Scalability

- The system supports scaling horizontally to accommodate increasing numbers of users, products, and data events.
- Storage and processing layers adapt to growing workloads without service interruption.

2.2.3 Availability & Reliability

- Core services target high availability (e.g., 99.9% or higher).
- Data durability is ensured through replication and backups.

2.2.4 Consistency & Correctness

- Data updates propagate consistently across components.
- Eventual consistency is acceptable in read-heavy scenarios, but transactions maintain correctness.

2.2.5 Security & Compliance

- Data in transit and at rest must be encrypted.
- Authentication and authorization mechanisms enforce role-based access.
- Sensitive data (e.g., PII) is protected through masking, retention limits, and deletion workflows.

2.2.6 Maintainability & Operability

- The system supports clear logging, monitoring, and documentation for operations teams.
- Infrastructure and deployments are automated and reproducible.

2.2.7 Cost Efficiency

- Data and compute resources are managed efficiently, using tiered storage and resource optimization.

3 User Stories

The user stories describes the main use cases from the user's perspective, including the related actions, the necessary conditions to do these actions, and the expected results of the actions.

Title: **Authentication**	Priority: **High**	Estimation: **10 h**
User Story: As a end-user or admin, I want to authenticate before the system; in case I don't have a user registered, I want to be able to create one. So that I can access the app resources and do various operations, according to my permissions level.		
Acceptance Criteria: • Given a valid username/email, when the user attempt login, then the app must allow to access the website. • Given a invalid username/email, when the user attempt login, the app must show an error message. • Given a valid form data, when a person attempt to register, the app must create a user. • Given a invalid form data, when a person attempt to register, the app must show an error message.		

Table 1: End User and Admin user story for authentication

Title: **Products Management**	Priority: **High**	Estimation: **10 h**
User Story: As a supplier user, I want to manage (create, edit, delete, and search) the products I desire to offer so that the end user can know them and buy them.		
Acceptance Criteria: • Given a valid form data, when the user attempt to create a product, then the product must be created, and the catalog updated. • Given a valid form data, when the user attempt to edit a product, then the product must be updated. • Given a confirmation, when the user attempt to delete a product, then the product must be deleted.		

Table 2: Supplier user story for products management

Title: **Custom Products Searching**	Priority: **High**	Estimation: **10 h**
User Story: As a user I want to do search by various criteria, so that the app throw me custom results based in my app interactions and history.		
Acceptance Criteria: Given a simple (by name) or complex input, when the user is searching a product, then the app must show me coincident results, based on my previous interactions and history, prioritizing brands and categories that I previously review or buy.		

Table 3: User story for custom search

Title: **Cart, checkout and order lifecycle**	Priority: **High**	Estimation: **10 h**
User Story: As a end user I want to buy and reserve products, so that I can acquire them.		
Acceptance Criteria: • Given a transaction, when the user is finalizing the buy process, then the payment method used by the user at the buy moment must be able to cover the transaction. • Given a valid transaction, when the user has finished the buy, then the inventory must be updated. .		

Table 4: End user story for buy products

Title: **Returns and refunds**	Priority: **High**	Estimation: **10 h**
User Story: As an end user I want to be able to request product refunds or changes, in case the acquisition don't fulfill my expectations.		
Acceptance Criteria: • Given a valid refund or product request, when the product don't fulfill user expectations, the app must do the referred request. • Given a valid refund or product request, when the product don't fulfill user expectations, the app must trigger a notification to the supplier, which have 3 calendar days to respond.		

Table 5: User story for refunds and returns

Title: **Business intelligence dashboards**	Priority: **mid**	Estimation: **32 h**
User Story: As a Admin, I want access to a management dashboard, so that the app facilitate me the business administration, viewing relevant metrics and user's preferences.		
Acceptance Criteria: Given a data request, when the user attempt to obtain info, then the dashboard must pull real time information.		

Table 6: User story for business dashboards

Title: **Profile Personalization**	Priority: **mid**	Estimation: **2 h**
User Story: As a user, I want personalize my profile, to obtain better recommendations.		
Acceptance Criteria: • Given a valid form data, when the user is updating his/her profile, the app must do an partial/total update. • Given a profile, when the user is searching, then the profile personalization must impact on the search recommendations.		

Table 7: User story for profile customization

4 DataBase Architecture

Presentation of the initial database architecture for the project

4.1 High-Level Architecture

Our high level architecture is based on two main parts a Date Lake that is going to use a No SQL technology (yet to be defined) that recollect information from:

- Web: Extracting information to simulate a steam of data.
- Behavioral patterns: Collect information from a variety of user actions.
- Partners: Supplier users in charge of providing data.

On the other hand, a Date Warehouse divided in 3 parts for different functionalities:

- Products: segment designed to manage the showing, filtering, and searching of products.
- Sales: main segment for the sale logic.
- User patterns: segment designed to manage the recommendation system.

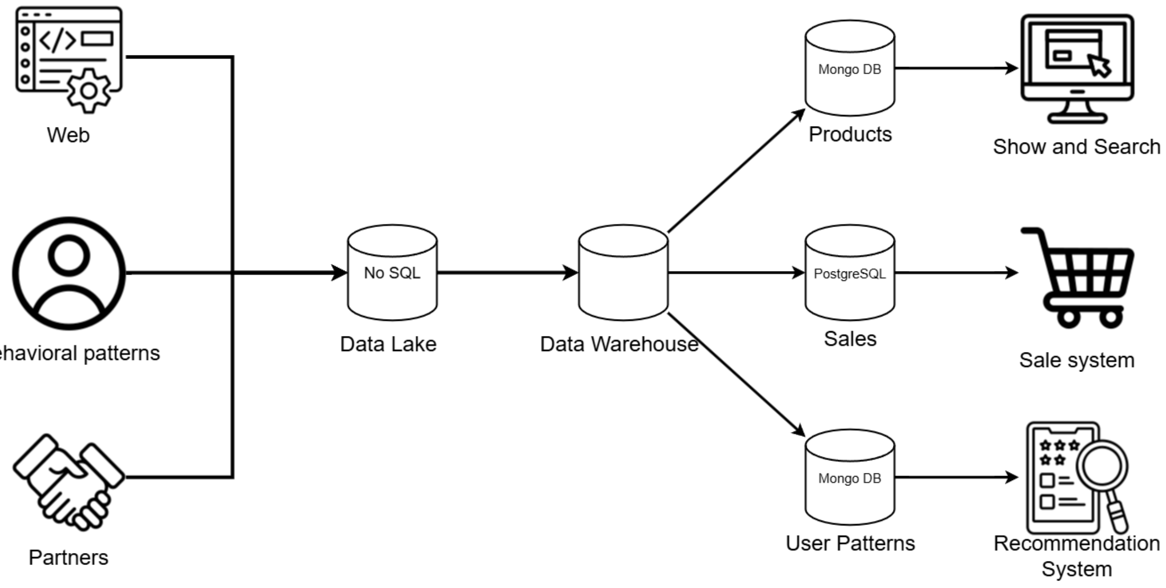


Figure 2: High Level DB Diagram

4.2 First version of ER Diagram

The core of the sale database are the entities user, product, and sale, which are related by the relationships "Offer" (a user offers a product), "Buy" (a user buys a sale), and the relationship between sale and product by the weak entity detail sale. Besides, sales have a state that can be updated.

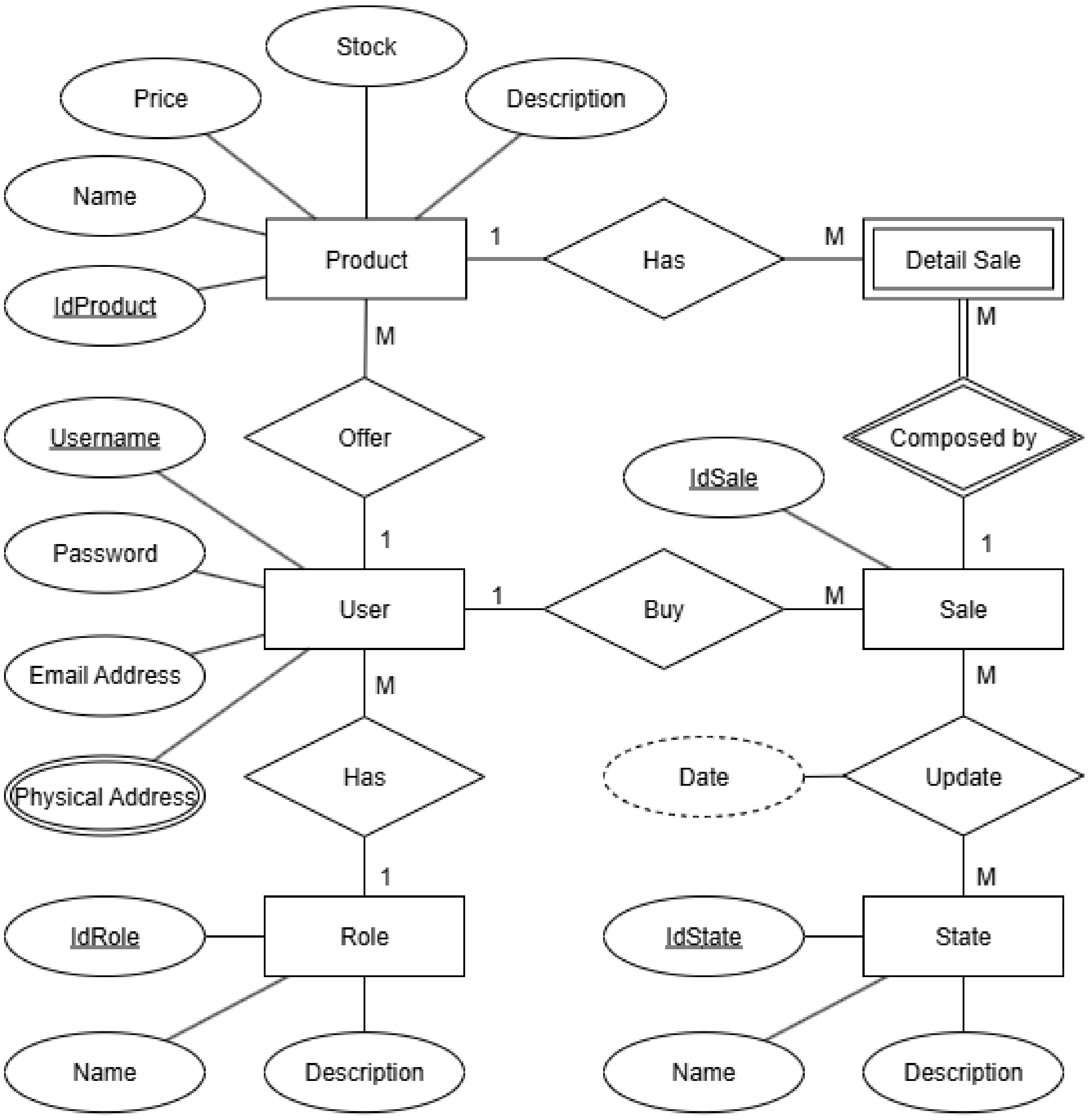


Figure 3: Entity-Relationship Diagram v.1.0

4.3 Storage Solutions

We propose a hybrid storage in order to reduce implementation's costs, potentially using a cloud storage solution for the data lake and an on-premise solution for a part of the data warehouse due to our familiarity with these alternatives.

5 References

References

[1] Osterwalder, A. & Pigneur, Y. Business Model Generation. Wiley, 2010.

[2] Kimball, R. & Ross, M. The Data Warehouse Toolkit. Wiley, 2013.

[3] Ricci, F., Rokach, L., Shapira, B. Recommender Systems Handbook. Springer, 2015.