Executive Report - CloudWalk Transactional Data Analysis and Strategic Decisions

Reference period: consolidated dataset (Q1 2025)
Data source: CSV files
Responsible for the analysis: Heloísa Zanati| Operations Intelligence Analyst

---

1. Context and Objectives of the Analysis

The purpose of this report is to provide a comprehensive and in-depth assessment of the operational and financial performance of the transactions processed by CloudWalk's payment system. More than just presenting numbers, this analysis seeks to interpret patterns, relate indicators, raise hypotheses and propose concrete actions that can result in financial gains, increased efficiency and better competitive positioning.

The data used in this analysis was extracted from .csv files, passed through python code and resulting in pre-processed and consolidated transactional data, which represents an accurate picture of the behavior of customers, products and payment methods over the period observed. The volume of records is significant (62,468 lines), covering tens of thousands of transactions and allowing consistent trends to be identified. In addition, the strategic dimensions that are already represented in the original Power BI dashboard were taken into account, allowing this document to function as a narrative complement to the BI tool, with the advantage of being able to circulate as an executive report between different areas of the company.

---

2. Overview and Volume Transacted

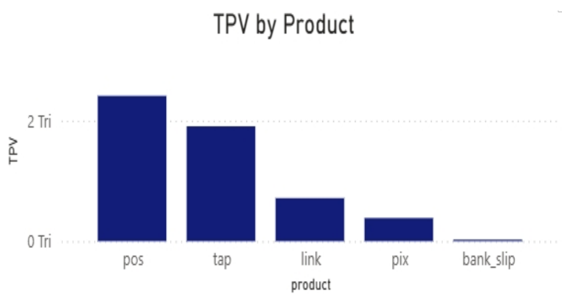
The central indicator of any payment operation is the **Total Payment Volume (TPV)**, which represents the sum of all the amounts processed in the period. In the set analyzed, global TPV reached the expressive mark of approximately **R$5.4 trillion**. This figure alone shows the magnitude of the operation, but the value becomes even more significant when broken down by entity segments, by product and by payment method.

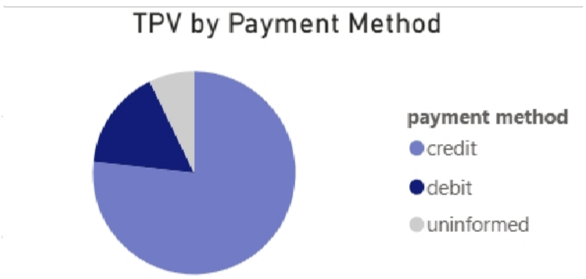
The analysis by entity reveals a clear predominance of Legal Entities (LEs), responsible for around two thirds of the total volume handled, amounting to approximately R$3.5 trillion, compared to R$1.9 trillion recorded by Individuals (IPs). This imbalance, far from being a problem, represents a clear opportunity: the PJ segment is the engine of revenue, and loyalty policies, portfolio expansion and the offer of higher value-added services for companies can generate substantial gains. At the same time, the PF segment should not be neglected, as it concentrates products and
tickets that can serve as an entry point for new customers, especially through digital channels.

When you look at TPV by product, there are still two main players: POS, with approximately R$2.41 trillion, and Tap, with R$1.9 trillion. Together, they account for almost 80% of total turnover. This predominance suggests that the

operation is heavily dependent on physical capture solutions, where user experience and equipment reliability play a key role in retention. Products such as Link (R$7.1 billion) and Pix (R$3.8 billion) appear to be
important alternatives, especially in niches where remote sales or instant payment are competitive differentiators. Boleto Bancário, with a residual share of R$68 million, remains relevant only in very specific segments or due to the demands of customers with a traditional profile.

The breakdown by payment method shows a significant dependence on credit cards, responsible for R$4.2 trillion, followed by debit (R$8.69 billion) and transactions classified as "uninformed" (R$3.91 billion). The dominance of credit indicates that installment payments are a central vector in sales dynamics and that any change in installment payment conditions can have a direct impact on total revenue. This
characteristic requires extra attention when formulating commercial policies, as adjustments to terms, rates and installment limits can be used as strategic levers to boost results.





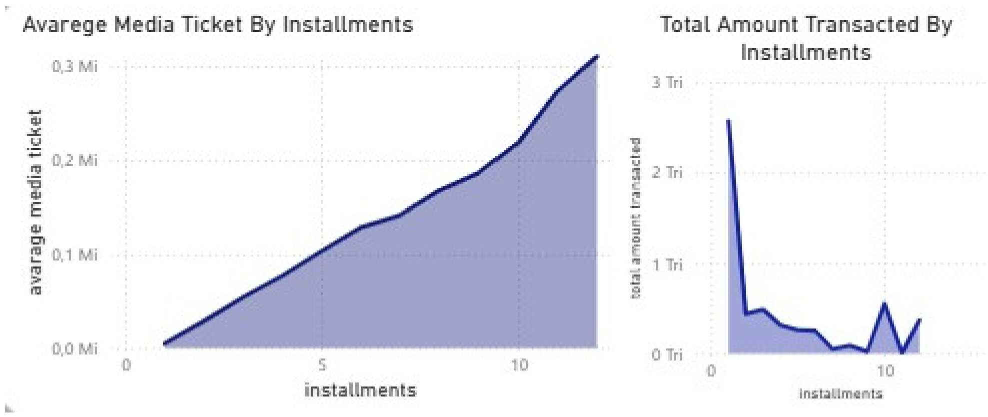
3. Average Ticket and Installment Dynamics

The average ticket, i.e. the average value per transaction, is one of the most important indicators in the market.
relevant to assessing the quality of revenue, at R$11.31 thousand. It reveals whether the growth in TPV comes from an increase in the number of transactions, an increase in unit values, or a combination of the two. In the base analyzed, the average ticket shows important variations between entities, products and payment methods.

Corporate clients tend to have higher average tickets, especially for products like Bank Slip (R$68,000) and Link (R$66,000). The same is true of the PF segment, but the values are more heterogeneous, with Link and Bank Slip (R$51,000) and Link (R$44,000), for cases classified as "high ticket". This reinforces the
the need for pricing and installment strategies adapted to the client's profile.

Statistical analysis indicates a **moderate positive correlation (close to 0.5)** between the number of installments and the average ticket, which means that, on average, the higher the number of installments, the higher the unit value of the transaction. This relationship is intuitive, since higher-value products tend to be paid for in more installments. However, when we look at the aggregate TPV by range of installments, the correlation is reversed, showing a negative value (-0.55) with the average ticket per installment. This suggests that sales with sales with many installments, although they have a higher ticket, do not necessarily represent significant volumes in the total.

This behavior indicates that installment policies need to be selective: encouraging the use of more installments for high value-added products can be positive, but generalizing this practice for all segments can lead to increased risk and unnecessary dilution of cash flow.

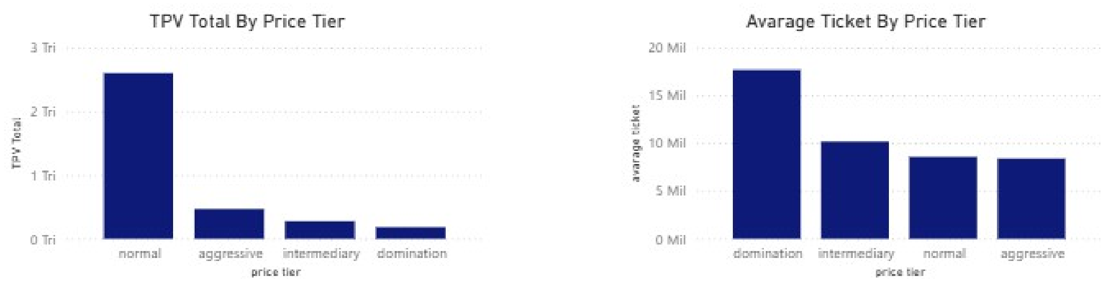


---

4. Price Tiers and Segmentation Strategy

Another relevant axis of analysis is the classification by price tiers, which group transactions according to unit value. The highest tiers, such as "Domination", have higher average values per transaction, but concentrate a lower volume of operations. The "Normal" tiers, on the other hand, account for a large part of the total TPV and a significant volume of transactions.

The strategic reading here is clear: high tiers demand differentiated treatment, with personalized service, exclusive conditions and retention strategies based on experience and relationships. Medium tiers, on the other hand, represent fertile ground for upselling campaigns, with incentives to increase the average ticket. On the other hand, lower tiers can be worked on with a focus on increasing volume and gradually introducing offers that encourage migration to more profitable ranges.

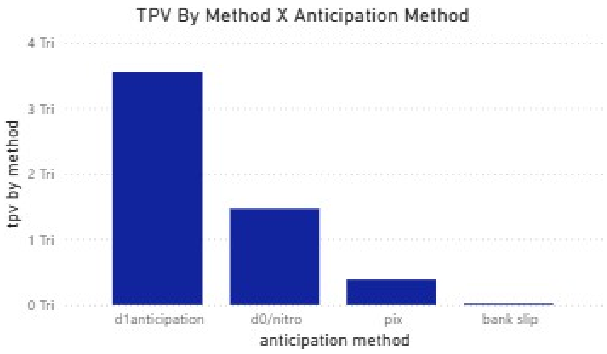


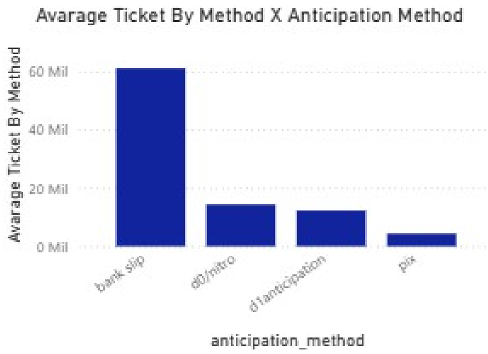
---

5. Anticipation Methods and Liquidity Management

Anticipation methods, such as D1, D0, aggressive or normal, not only affect the company's cash flow, but also influence the attractiveness of the solution for different customer profiles. In the set analyzed, it can be seen that faster prepayment methods tend to concentrate higher average tickets, which suggests that clients
willing to pay for the advance of the amounts transacted are, in general, those who process higher value transactions.

This relationship is strategic: by encouraging more profitable methods of anticipation, the company can not only increase net margin, but also build loyalty among high-value customers by offering them tangible benefits such as rate reductions or customized conditions. On the other hand, for lower-value segments, immediate prepayment can be offered as an additional service, monetizing an audience that, although not high-ticket, values immediate liquidity.



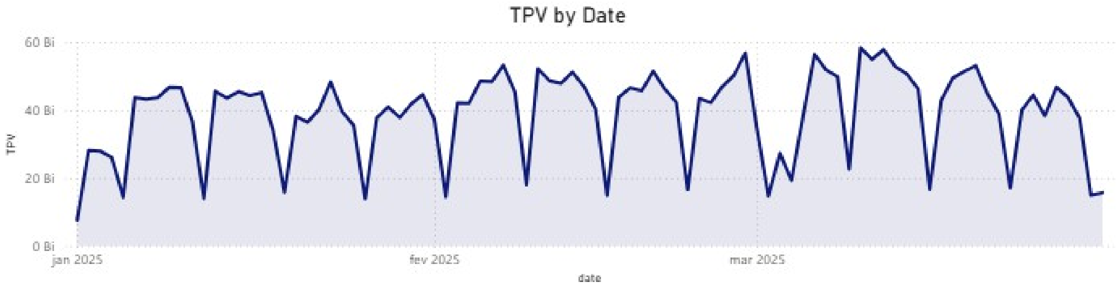


---

6. Seasonality and Time Patterns

Analysis of transactions over time shows that volumes are concentrated on certain days of the week, especially Wednesdays and Fridays.
days of the week, especially Wednesdays and Fridays. This suggests that there is a cyclical behavior that may be related to customers' commercial practices, salary dates, promotion calendars or even the consumption habits of the end public.

Identifying these patterns is fundamental to aligning resources and strategies. On peak days, it is advisable to reinforce support teams, guarantee robust systems and even program commercial campaigns that take advantage of the increase in traffic. A
he absence of hourly data prevents a more granular analysis, but it is strongly recommended that this information be incorporated into future analyses, making it possible to identify not only the days, but also the times of greatest demand.



weekday	Média de ticket_medio	Média de quantity_transactions
Friday	28473067817,06	10420,51
Monday	28511180010,89	8197,31
Saturday	28471967621,27	12407,45
Sunday	26108328650,65	9563,62
Thursday	28365901942,49	9454,57
Tuesday	28567979385,59	8823,33
Wednesday	28737987358,82	8870,73
Total	**28248818187,12**	**9686,88**

---

7. Conclusion and Next Steps

Conclusion

Analysis of CloudWalk's transactional data in Q1 2025 reveals a robust operation concentrated on large volumes, with a TPV of R$5.4 trillion. Legal Entities are the main driver of revenue, concentrating two thirds of the volume
volume, while POS and Tap products stand out as the predominant capture channels, indicating the strength of physical solutions. The credit card remains the dominant method, highlighting the importance of installments in commercial dynamics - especially in higher average ticket transactions, where there is a positive correlation with the number of installments. However, bands with many installments do not proportionally represent the largest aggregate volumes, which requires a selective approach to financing policies.

In addition, segmentation by price tiers and anticipation methods offer clear opportunities for strategic action: high-value customers demand exclusive offers and tailor-made conditions, while rapid anticipation is related to higher tickets and can be a lever for loyalty and profitability. Finally, seasonal
seasonal patterns - especially Wednesdays and Fridays - indicate ideal windows for promotional actions and operational reinforcement. The future incorporation of hourly data and continuous analysis of these dimensions will make it possible to further refine strategies for growth, efficiency and customer relations.

8. Application of Automation

8.1 Continuous Transactional Data Ingestion

How to Apply

Continuous data ingestion is essential to ensure that transactional information is always up-to-date and ready for analysis. The system must be able to capture data in real time or at regular intervals, processing it to ensure quality and integrity.

8.1.2 Implementation:

8.1.2.1 Ingestion Pipeline Architecture:

The ingestion pipeline architecture can be divided into four main stages: Extraction, Validation, Transformation and Load (ETL). Below are the details of each of these stages.

a. Extraction

- Data Sources: Data can be extracted from various sources, such as:
- AWS S3: Storage of CSV files with transactional data.
- APIs: RESTful endpoints that provide real-time data.
- Webhooks: Real-time notifications of transaction events.
- Frequency: Extraction should be scheduled to take place daily, preferably at 00:30, just after the previous day's close.

b. Validation

- Integrity check: Ensure that the extracted data meets quality criteria, such as:
- Presence of mandatory columns (e.g. `amount_transacted`, `quantity_transactions`).
- Correct data types (e.g. `amount_transacted` must be numeric).
- Missing or anomalous values (e.g. `amount_transacted` equals 0).
- Validation example in Python:

```python
def validate_data(df):
  required_columns = ['amount_transacted', 'quantity_transactions', 'daydia']
  for column in required_columns:
    if column not in df.columns:
      raise ValueError(f"Column {column} not found.") if
  df['amount_transacted'].isnull().any():
    raise ValueError("Missing values found in 'amount_transacted'.")
```

c. Transformation

- Data enrichment: Apply business logic to calculate KPIs and derive new fields. Examples include:
- Calculating the average ticket.
- Classifying transactions into price tiers.
- Deriving temporal fields (e.g. day of the week, month).
- Transformation example in Python:

```python
```

```python
def transform_data(df):
    df['ticket_medio'] = df['amount_transacted'] / df['quantity_transactions']
    df['is_high_ticket']= df['ticket_medio']> df['ticket_medio'].quantile(0.9)
    df['day_of_week']= pd.to_datetime(df['daydia']).dt.day_name()
    return df
```

d. Load

- Database Persistence: The transformed data must be loaded into an analytical database, such as Amazon Redshift or PostgreSQL, for efficient queries.
- Example of loading in Python:

```python
from sqlalchemy import create_engine


def load_data(df):
    engine = create_engine('postgresql://user:password@host:port/dbname')
    df.to_sql('transactions', engine, if_exists='replace', index=False)
```

8.1.2.2 Orchestration with Apache Airflow

Apache Airflow can be used to orchestrate the ETL pipeline, allowing tasks to be scheduled and monitored.

- Example of a DAG in Airflow:

```python
from airflow import DAG
from airflow.operators.python_operator import
PythonOperator from datetime import datetime


def etl_process():
    raw_data = extract_data()
    validate_data(raw_data)
    transformed_data= transform_data(raw_data)
    load_data(transformed_data)
```

```
dag= DAG('etl_pipeline', start_date=datetime(2023, 1, 1), schedule_interval='@daily')


etl_task = PythonOperator(
  task_id='run_etl',
  python_callable=etl_process,
  dag=dag
)
```

8.1.3 Expected Results:

- Updated data: Continuous ingestion ensures that data is always up-to-date, allowing for real-time analysis.
- Data Quality: Continuous validation reduces errors and inconsistencies, increasing confidence in analysis.
- Operational Efficiency: Automating the ingestion process reduces the time and effort required to collect and process data manually.

8.1.4 Required Data:

- CSV Structure: CSV files must contain columns such as: `amount_transacted`, `quantity_transactions`, `day`, `entity`, `product`, `payment_method`, among others.
- Access to APIs: If the extraction is done via API, it is necessary to have the API credentials and documentation.

8.2 Anomaly and Pattern Detection

How to apply it

Anomaly detection is important for identifying unexpected behavior in the data, such as sudden drops in TPV or spikes in cancellations. The system must be able to continuously monitor KPIs and trigger alerts when anomalous patterns are detected.

8.2.2 Implementation

8.2.2.1 Definition of anomalies:

- Drop in TPV: If the daily TPV drops by more than 15% in relation to the 7-day moving average.

- Spikes in Cancellations: If the cancellation rate exceeds the average+ 3 standard deviations.
- Changes in Price Tiers Mix: If there is significant volume migration between price tiers.

8.2.2.2 Rules Engine:

Anomaly detection rules can be defined in a readable format such as JSON and implemented in a Python script that evaluates these rules at regular intervals.

- Example of Rules in JSON:

```json
{
  "rules": [
    {
      "name": "drop_on_tpv",
      "condition": "tpv_today < 0.85 * moving_avg_7d",
      "action": "alert",
      "severity": "high",
      "description": "TPV has fallen by more than 15% compared to the 7-day moving average"
    },
    {
      "name": "peak_cancellations",
      "condition": "cancel_rate> mean+ 3 * std_dev",
      "action": "alert",
      "severity": "critical",
      "description": "Cancellation rate above critical threshold"
    }
  ]
}
```

8.2.2.3 Implementation in Python:

- Example Code for Anomaly Detection:

```python
```

```python
import pandas as pd


def detect_anomalies(df):
    df['moving_avg'] = df['amount_transacted'].rolling(window=7).mean()
    df['threshold'] = df['moving_avg'] * 0.85 # 15% decrease
    anomalies= df[df['amount_transacted']< df['threshold']]
    return anomalies


# Usage
transactions_df= pd.read_sql('SELECT * FROM transactions', engine)
anomalies = detect_anomalies(transactions_df)
if not anomalies.empty:
    print("Detected anomalies:", anomalies)
```

8.2.3 Expected results:

- Rapid identification of problems: detecting anomalies allows the team to react quickly to operational problems, minimizing financial losses.
- Improved Decision Making: With real-time alerts, managers can make informed and proactive decisions.
- Increased Reliability: The ability to detect anomalies increases confidence in analyses and reports.

8.2.4 Types of Alert Messages:

- TPV Drop Alert [ALERT]

TPV Drop Detected

Description: TPV fell by 18% in relation to the 7-day moving average.

Recommended Action: Check for possible payment gateway problems or changes to marketing campaigns.

- Spikes in Cancellations Alert:

[CRITICAL ALERT] Increase in Cancellation Rate

Description: The cancellation rate has exceeded 5%, 3 standard deviations above the average.

Recommended Action: Investigate possible fraud or user experience issues.

8.2.5 Required Data

- Historical Transactional Data: To calculate moving averages and standard deviations.
- Clear Rule Definitions: Anomaly detection rules must be well defined and documented.

8.3 Generating Actionable

Recommendations How to Apply

Generating actionable recommendations allows stakeholders to make informed decisions based on data. The system should analyze KPIs and suggest specific actions to optimize performance.

8.3.2 Implementation

8.3.2.1 Recommendation Logic:

The recommendation logic should be based on heuristics and machine learning, analyzing trends and patterns in the data.

- Example code in Python
- Example of Recommendation Logic:

```python
def generate_recommendation(tpv_trend):
    if tpv_trend < -0.15:
        return "Check gateway and active campaigns."
    elif tpv_trend > 0.1:
        return "Consider increasing the reach of campaigns." else:
        return "Monitor stability."


# Use
tpv_trend= (current_tpv - previous_tpv) / previous_tpv
recommendation = generate_recommendation(tpv_trend)
print("Recommendation:", recommendation)
```

- 8.3.2.3 Bot integration

The conversational assistant can be integrated to provide recommendations in real time, allowing users to ask questions and receive contextual answers.

8.3.3 Expected Results

Increased Operational Efficiency: Data-driven actions help optimize processes and improve financial results.

Informed Decisions: Stakeholders have access to clear, actionable recommendations, facilitating decision-making.

Reduced Analysis Time: Automating the generation of recommendations reduces the time spent on manual analysis.

8.3.4 Types of Recommendation Messages

- Recommendation for Action in the Event of a Drop

in TPV: [RECOMMENDATION] Action Required

Description: TPV fell by 20% compared to the previous day.

Recommended Action: Check gateway performance and consider a promotional campaign to reverse the drop.

- Recommendation for Action in Case of Growth in TPV:

RECOMMENDATION] Growth Opportunity

Description: TPV increased by 15% compared to the weekly average.

Recommended Action: Increase investment in marketing campaigns and consider product expansion.

8.3.5 Required Data

- Calculated KPIs: Data such as TPV, average ticket and cancellation rates are essential for generating recommendations.
- Action History: Information about previous actions and their results helps refine recommendations.

These three ideas form the basis of a robust system that not only monitors performance, but also provides valuable insights for optimizing operations and increasing efficiency.