**Phase 3 Report: Ardian PE Vespucio Project**

Michelle (Meixi) Peng, Vidhi Agrawal, Kris (Tiantian) Chen, Raphael Barthes, Howard (Haozhou) Yu

1. **Background Summarization**

The Vespucio project, a landmark infrastructure initiative in Santiago, Chile, represents a significant advancement in urban toll road systems. Orchestrated in partnership with Ardian PE and involving local investor CMB Prime, this project focuses on two major toll road concessions: the Autopista Vespucio Norte (AVN) and the Túnel San Cristobal (TSC). These concessions, integral to Santiago's transportation network, were acquired by Ardian with an investment reflecting a net asset value of $183 million.

1. Understanding Concessions and Free-Flow Tolling

Before delving into the complexities of the project, it is essential to clarify some key terms. In the context of the Vespucio project, a "concession" refers to a public-private partnership agreement. Under this arrangement, a private entity, like Ardian, is granted the rights to operate and maintain a public infrastructure, such as a toll road, for a specified period. In return, the concessionaire collects tolls, generating revenue to offset their investment and operational costs.

The Vespucio project utilizes an innovative toll collection method known as "free-flow tolling." This system, unlike traditional toll booths, allows for the electronic collection of tolls. Vehicles equipped with electronic tags or license plate recognition technology can pass through toll areas at highway speeds, enhancing traffic flow and reducing congestion. This method, while efficient, presents unique challenges in terms of toll evasion, commonly referred to as "leakage."

2. Importance to Ardian

Ardian holds ambitious investments in the Americas. Having successfully extended its infrastructure strategy from Europe to the Americas, the firm focuses on essential infrastructure diversified across various verticals and aims to create value through an industrial mindset, energy transition, and data-driven infrastructure.

The Vespucio project is an essential part of this strategy. These toll roads are not only crucial transportation links in Santiago, connecting key areas including the international airport, but also represent Ardian's commitment to sustainable and efficient infrastructure. Its alignment with Ardian’s values ensures its integral role in the firm's portfolio, and its successful implementation is expected to demonstrate Ardian's capability to leverage technology to enhance infrastructure efficiency and profitability.

The project also bolsters Ardian's financial strategy by offering stable and predictable revenue streams, crucial for long-term investment planning. The concessions, due to their strategic locations and consistent traffic flow, provide a dependable source of income, with contracts offering visibility into cash flows for an average of 12 years plus potential extensions. Furthermore, the project's tariff structure, indexed to the Chilean CPI and including congestion charges, ensures revenue keeps pace with inflation and adapts to traffic patterns. This combination of stable revenue, long-term financial visibility, and a dynamic pricing strategy aligns closely with Ardian's goals, enhancing its portfolio's resilience and growth potential.

3. Challenges and Objectives

The primary challenge in the Vespucio project arises from the inherent nature of free-flow tolling, which depends largely on the honor system and is susceptible to toll evasion. This issue is significant as it impacts toll collection efficiency – a vital factor for the project's financial health. Inefficiencies in collecting tolls directly influence the project's bottom line, emphasizing the necessity for enhanced efficiency in toll collection to maximize revenue and ensure the project's long-term sustainability.

Our project's central objective is to tackle this challenge, aiming to refine toll collection efficiency, which is essential to reduce revenue leakage and guarantee the financial viability of the concessions. This goal will be achieved by harnessing the extensive data at our disposal to pinpoint and address inefficiencies within the current toll collection system. Our approach will involve a thorough analysis of the data, through which we will gain insight into the project’s current problem and operational processes, and develop strategies to mitigate revenue loss. In the final stages of our project, we will integrate our solutions with Vespucio’s existing operational strategies, and analyze the impact of our suggestions on Vespucio’s toll collection rate, bottom-line, and other financially significant factors.

1. **Data Description**

Our team will primarily focus on analyzing the Vespucio dataset, which comes from the Vespucio data project, an implementation of a cloud data warehouse leveraging Snowflake Data Cloud. Advised by Ardian’s IT and Data Science team, Vespuccio is integrating isolated data silos, which allows for improved management and operations. The project has successfully onboarded 470 million transactions with its automated data processing and updates, organizing and storing them in 59 tables on Snowflake. From the dataset dictionary we have been introduced, this dataset contains detailed information on customer characteristics, behaviors, and payment data. For our analysis, we will potentially also incorporate relevant macroeconomic indicators, such as Chilean inflation data, to ensure a comprehensive understanding of external influences on our objective.

Since access to the Vespucio database has not yet been granted and we have only been given snippets of the data, only limited information has been collected. 5 out of 59 tables are provided as samples to explore example data elements and structures. These 5 tables encompass aspects of toll road operations, including toll rates, point collection, vehicle categories, billing cycles, and potential income. In order to interpret categorical/acronymed columns, we have reached out to request dataset descriptions for these tables, but have not yet received a response. Regardless, each table can still offer us some insights. Below are brief descriptions of the 5 provided tables:

1. Toll Rates

This table contains information about the toll rates, including extraction date, the applicable year, a toll point code, vehicle category, day type, the monetary value of the toll, start and end times, tariff information, concession, and direction. This dataset is fundamental in understanding how toll rates are structured and vary with toll point, time of the day, vehicle type, and direction, which are critical factors in revenue calculation and optimization.

1. Point Collection

This table focuses on the operational aspect of toll collection, including detailed information about each toll collection point, such as booth number, domain ID, travel direction, location in kilometer point, description, previous toll booth, section, and whether it's active. It is crucial for understanding traffic flow and identifying potential bottlenecks or inefficiencies in toll collection.

1. Vehicle Category

This table classifies vehicles based on their category, alongside the extraction date and account contract numbers. Vehicle categories are pivotal in toll calculation as different rates apply to different categories. Understanding the distribution of vehicle types using the toll road can aid in tailoring toll collection strategies to maximize revenue from each category.

1. Billing Cycles

This table captures the billing cycle in terms of the process timeline. It includes the extraction date, specific billing cycle identifiers, the dates marking the beginning and end of a billing period, the accounting date when the billing records are finalized, and a path to access related documents. Analyzing this data helps track the efficiency and timeliness of the billing cycle and identifies delays or inconsistencies that might affect revenue collection.

1. Potential Income

This table estimates potential revenue from tolls, listing the dates of transactions, concessionaires, toll booth numbers, vehicle categories, and specific details like day, hour, time slot, and direction of travel. It also specifies the tariff type and the period it applies to, alongside the quantity of transactions and the total amount expected. This dataset is essential for forecasting revenue and identifying patterns, which could inform dynamic strategies to optimize income.

There are common columns that we could use to join tables and analyze inter-relational patterns, such as toll booth number, vehicle category, extraction date, and direction. There are also similar value columns, such as TariffType in Potential Income and Tariff\_Low, Initial\_Tariff, and Tariff\_Period in Rates Toll tables, suggesting a connection between specific tariffs and revenue outcomes.

The initial analysis of the Vespucio dataset provides a crucial baseline understanding of toll operations and revenue potential, despite being limited to a fraction of the complete dataset. This groundwork prepares us for a deeper dive into data-driven optimization strategies once full dataset access is granted. Our next steps will include employing clustering techniques to refine our understanding and enhance toll collection efficiency, setting the stage for achieving financial sustainability through informed decision-making and strategic optimizations.

1. **Timeline (Subject to Review)**

Our current expectation of the project timeline is shown as follows. It should be noted that there is not yet a guaranteed date on which we will gain access to the database, and the specific schedule of Ardian’s team in France (to which we are expecting to present our final report) is yet to be determined. Therefore, the timeline below will be revised in light of future information.

| Date | Content | Notes |
| --- | --- | --- |
| 2/12 - 2/18 | Cleaning of the limited data, translation (Spanish-English), discover correlation among variables, and consolidate all provided datasets.  Research and literature review on loan defaults, payment collection and customer clustering. | It is expected that Vespucio will provide us with the complete dataset by 2/22, on which their staff will walk us through the data |
| 2/19 - 3/10 | Exploration of the complete dataset, finalization of the clustering and segmentation of customers based on behavioral data, and construction of an initial collection strategy. | Weekly presentation to Ardian on data exploration (15 - 20 minutes) |
| 3/18 - 3/31 | Completion of the collection strategy. | At the end of March, we are expected to conduct a Mid-term presentation to Ardian. |
| 4/1 - 4/14 | Further improvement and revision of our proposed collection strategy, and begin working on our final presentation to Ardian. | Final Presentation should include quantified impacts of our suggested strategy, as well as how the solution can be integrated with Vespucio’s existing operations. |
| 4/15 - 4/28 | Carry out the final presentation to Ardian. |  |