**Case Study – AdventureWorks Cycles**

AdventureWorks sells bicycles and bicycle parts directly to customers and distributors. The company currently has a single office in the Netherlands, and have been selling bicycles in the United States, Germany and Spain through a chain of distributors and through online sales on its website. The fulfillment of delivery is done by local distribution centers.

The company is planning to expand by establishing new offices because the sales growth in these countries has been increasing over the last 3 years. The location are:

* Tokyo, Japan
* Seattle, USA
* Chicago, USA
* Berlin, Germany
* Barcelona, Spain
* Paris, France

In a highly competitive market, in which AdventureWorks has been in business for the last 15 years, it wants to become the most innovative bicycle company, providing both current and future bicycle owners with best in class technology and service that provides unique experiences.

The Research and Development department of AdventureWorks has successfully conceived the next wave of innovative products, and they are relying on Data Engineers, AI Engineers and Data Scientists to assist with both the design and implementation of the solution.

Given the increased level of sales and expansion at global scale, the existing data infrastructure won't meet the overall business requirements or the future growth that AdventureWorks aspires to. The Chief Information and Technology Officers have expressed the desire to abandon existing on-premises systems and move to the cloud to meet the growth expected. This is supported by the CFO as there has been a request for replacement hardware as the existing infrastructure comes to its end of life. The CFO is aware that the cloud could offer alternatives that are more cost efficient.

As a Senior Data Engineer, you will assist AdventureWorks in the solution design and implementation to meet the business, functional and technical requirements that the company has set forth to be successful for growth, expansion, and innovation strategies. You will execute this in a way that minimizes operational costs and can be monitored for effectiveness.

In a discovery workshop you ascertained the following information:

**AdventureWorks Website**

The web developers at AdventureWorks are transferring the existing website from an on-premises instance of IIS, to an Azure Web App. They have requested that a data store is made available that will hold the images of the products that are sold on the website.

**Current Sales / Ordering system**

The current software on which bicycle purchases are tracked, is a web-based application which directly stores order information into an on-premises SQL Server database named AdventureWorks2012. The current application is deployed with high-availability provided by SQL Server 2012 Always-on Availability groups. Due to global expansion and data governance requirements, AdventureWorks will transition this system to better serve their customers and will be looking for global availability of its application and data sales and ordering purposes, particularly during the months of November and December when demand for bikes grow ahead of the holiday period.

**Data Analysis**

The business reporting is currently being provided by a single on-premises database that is configured as a data warehouse, it holds a database named AdventureWorksDW which is used to provide historical reporting and descriptive analytics. In recent times, that server has been struggling to process the reporting data in a timely manner, as a result the organization has evaluated the data warehouse capabilities of Azure SQL Data Warehouse and want to migrate their on-premises data to this platform. Your team should ensure that access to the data is restricted.

In addition, AdventureWorks would like to take their data analytics further and start to utilize predictive analytics capabilities. This is currently not an activity that is undertaken. The organization understands that a recommendation or a text analytics engine could be built and would like you to direct them on what would be the best technology and approach to take in implementing such a solution that is also resilient and performant.

**Customer Service / Presales**

Customer service and pre-sales departments are currently experiencing scale issues due to the high call volumes. The organization wants to support the customer services staff in handling the call volumes through the implementation of chat bots in which future bicycle owners can:

* Find which bicycle is best for them:
  + Through a set of questions with the chat bot, custom recommendations are given to potential bike owners, who then can take the recommendation and place an order, or can be redirect to a sales specialist to help them with their needs
* Check status on current orders:
  + Retrieve status on current orders, and estimated delivery times
* Find bicycle parts suitable for their existing bicycle:
  + Existing bicycle owners can find recommended bicycle parts and accessories based on the serial number or model number of their bicycle
  + Existing bicycle owners, can upload a picture of their bicycle or take a picture of the serial number of their bicycle to assist with the identification of their bicycle and have recommended bicycle parts

Given its global expansion, the customer service / presales chat bot needs to respond to requests for data in near real-time regardless of where the customer is located. The chatbot should also support multiple languages such as Dutch, German, French, English, Spanish, and Japanese. This work will be handled by the AI Engineers, but they have requested a platform is provided by the Data Engineer that enables them to store conversation history.

**Social Media Analysis**

In recent years, the marketing department at the organization have run a wide variety of twitter campaigns at various times of the year. They are keen to measure the impact of their work by tracking social media assets such as hashtags during those campaigns. They would like to have the capability of tracking any hashtag of any name.

**Connected bicycle**

AdventureWorks Bicycles can be equipped with an innovate built-in bicycle computer which consist of automatic locking features of the bicycle, as well as operational status. Information captured by this bicycle computer includes:

* Bicycle model, serial number and registered owner
* Bicycle location (latitude longitude)
* Current status (stationary, in motion)
* Current speed in kilometers per hours
* Bicycle Locked / Unlocked
* Bicycle parts and components information (on electrical bicycles)

First party and 3rd party applications can have access the information of the bicycle computer that must be secure and for the integration into mobile applications and real time display of location and bike ride sharing information.

Furthermore, daily summary data can be saved to flat files that include Bicycle model, serial number, registered owner and a summary of the total miles cycled per day and the average speed.

**Bicycle Maintenance services**

Existing bicycle owners can opt in to getting notifications on when their bicycle needs repair, based on:

* Telemetry from electrical bicycle based on sensor data
* Bicycle usage information coming from the built-in bicycle computers based on average mileage / wear and tear

This predictive maintenance scenario is a service in which bike owners can opt-in, offered as a paid service.

Finally, all services that are proposed should have a comprehensive business continuity that meets the corporate objective of minimizes restore times when recovering the data for a given service.