**Answer #1:**

This phase ensures that I gather the correct data sources, evaluate data quality, and identify key data elements that will inform my decisions.

Chosen Data and Rationale

The specific business problem focuses on improving operational efficiency and customer retention, and the data chosen reflects this. For this purpose, I will need a mix of structured and unstructured data, including:

*Customer Data:* Historical customer behavior, transaction records, and feedback. This data helps us understand customer needs, shopping patterns, and pain points, which is essential for optimizing customer retention strategies.

*Operational Data*: Data from internal operations, including production logs, inventory management, and supply chain efficiency. Walmart’s vast supply chain is critical to its business model, and this data will help identify bottlenecks and inefficiencies.

*Sales and Financial Data*: Sales history, revenue reports, and profit margins are essential to measure financial health and identify trends across Walmart's many stores and e-commerce channels.

*Employee Performance Data*: To understand internal workflow bottlenecks, it will be important to include data related to employee performance metrics, task completion rates, and workload, especially in store operations and logistics.

These datasets were selected because they directly align with Walmart’s key performance indicators (KPIs), such as customer satisfaction, operational efficiency, and profitability, and will be pivotal in measuring the effectiveness of the Agile analytics program.

**Answer #2**

Possible Data Sources

To gather this data, the following sources will be utilized:

*Internal Databases*: Walmart’s existing databases will be the primary source for operational, sales, and customer data. This includes CRM systems, ERP systems, and historical sales data repositories.

*Publicly Available Data*: External data from government databases, industry benchmarks, or publicly available research reports will provide additional context and competitive comparisons within the retail industry.

*Third-Party Vendors*: Data providers such as Nielsen, Gartner, or retail industry platforms can provide market insights and benchmarks to compare Walmart’s performance against industry standards.

*Surveys and Customer Feedback Systems*: Data directly from Walmart’s customer surveys or feedback systems (e.g., post-purchase reviews) will provide qualitative insights that support the structured data.

**Answer #3**

Key Data Elements:

*Customer Segmentation Data*: Demographics, purchasing patterns, and lifetime value are key to understanding which customer segments are most profitable and where churn may occur. This will help Walmart tailor its promotions and customer experience strategies.

*Product/Service Performance Data*: Metrics such as product return rates, defect logs, or service feedback will help pinpoint areas where improvements are needed in Walmart's product offerings or customer service.

*Process Efficiency Metrics*: Data regarding lead times, task completion rates, and supply chain metrics will help identify inefficiencies in Walmart’s operations, especially in warehouse and distribution center management.

*Financial Metrics*: Revenue growth rates, profit margins, and cost structures are necessary to measure the financial impact of the proposed solutions, particularly in optimizing store operations and reducing operational costs.

Each of these elements was selected because they directly link to solving Walmart’s business problem and measuring the success of the analytics program. For example, improving customer retention can be tied to behavioral data, and operational improvements can be linked to process efficiency metrics, which are critical in Walmart’s high-volume, data-intensive environment.