**This activity is continuing from the previous activity.**

|  |  |
| --- | --- |
| **Your task.** | In this activity, you need to define data warehouse steps and process so it will be easy to design and implement data warehouse for the Devon Technical college.  In the report, you need to complete the following information:  •       Overview of the report  •       Identify and list data warehouse information sources based on existing database design documents  •       Based on business process, data and policies, define data warehouse targets  •       Based on system configuration and requirements, identify two warehouse agents  •       Define data warehouse implementation steps  •       Prepare two processes related to data warehouse  •       Prepare data warehouse user interface  o   Prepare data table for each stakeholder  o   Identify primary and foreign keys  o   Link the table  o   Identify the relationship  o   Identify dimension and fact tables  •       Prepare warehouse data security strategy  o   Identify who is accountable for security  o   How to test the warehouse to identify security issues  o   How to manage the security patches  o   Where to store all the security issues  •       Conclusion |
| **Resources required** | Your trainer will provide you the following resources to complete this task:  •       Software to create warehouse user interface  •       MS office  •       Computer  •       Internet  •       Appendixes |
| **You will be required to complete.** | Your trainer will observe your performance and complete the following performance criteria checklist. |

Write answers to the following template.  Copy and paste to online text.

Student Pinyin name :ChenShanshan

Student English name:Shine

|  |
| --- |
| **Devon warehouse operational steps, processes and data warehouse catalogue** |
| |  |  | | --- | --- | | **ORGANIS** | **Devon Technical College** | |
| **Overview of the report**  **Devon Institute of Technology is a private registered training institution (RTO) located in the west of Melbourne. It covers many subject areas. Need to design Devon Warehouse and Data Warehouse for his school.** **The data warehouse synchronizes the work of all departments. It looks at all aspects of the college, students, departments, departments, grades and other extracurricular activities.** |
| **Identify and list data warehouse information sources based on existing database design documents**  **The key information source is ID**  **ID is important for DTC data warehouse and can be used as clustered index to improve query efficiency. If you add a unique user ID to a feature, then all the other features are unique to that user. There are four kinds of stakeholders: students, teachers, staff and visitors. The methods of information collection are different. Students and teachers search by their IDS, while others can only search part of Devon Technical College.** |
| **Based on business process, data and policies, define data warehouse targets**  **The primary goal of a data warehouse is to provide end users with a unified, flexible and meaningful data store for reporting and analysis. DTC The advantage of this system is an integrated Web application for handling the university's various academic and non-academic activities. The system is accessible to every student/department/employee of the institution via an Internet-connected computer. The system provides a detailed structure of the university campus and its departments. It can also link all the various departments of the organization, such as management, attendance, employee details and many other professional modules. Can timely access to the data, "higher data quality, reliability and low or zero redundant", "to achieve outstanding business decisions", "lower time consumption", "to store or archive historical data for later use", the data warehouse is dedicated to handle a large number of date and numerous complex queries, so it is of high functional core DTC data analysis practice. Efficient data analysis of DTC.** |
| **Based on system configuration and requirements, identify two warehouse agents**  **The warehouse agent is the component of the DB2 Warehouse Manager that manages the flow of data between data sources and targets that are on different computers. The computer on which you install a warehouse agent is called an agent site. When the ETL is running, the control server instructs the remote warehouse agent to run the SQL commands associated with the ETL. The agent uses the database client of the source database to run commands on the source or target database. As data is transmitted, it passes through the proxy site. To reduce network traffic to control servers, carefully plan where to place proxy sites in your environment and the sites associated with each ETL. The control server uses the local warehouse agent to manage the flow of data between the operational data store, the central data warehouse, and the data mart. Proxy sites can also be created on other Windows and UNIX systems，as well as SQL Server agents and Operations Manager agents.** |
| **Define data warehouse implementation steps**  **1．Identify and collect requirements：We want DTC to interview key decision makers to understand what factors determine the success of DTC business. We also need to cooperate with people from different departments to understand the data and their common relationship (if any) and record their entire needs, which needs the system to meet.**  **2. Data analysis：Conduct data analysis on DTC schools, understand the information of each student and teacher, import data, and connect all departments of the institution, such as management, attendance, employee details and many other professional modules.**  **3. Business actions：Create and fill dimension tables and fact tables, and the insights and information obtained from data analysis will be further used in the organization's decision-making.** |
| **Prepare two processes related to data warehouse** |
| *Process 1* how to use data warehouse |
| **The data warehouse stores the data, and DTC School's smart platform analyzes it. When you make these two systems work seamlessly together, you unlock the full benefits of DTC school's smart platform. DTC should also consider using business applications to connect to SQL servers.**  **Data extraction: Obtaining data from the source system for subsequent data warehouse use. The choice of extraction method is highly dependent on the DTC business needs of the source system and target data warehouse environment.**  **Data conversion: After extraction, data needs to be transformed, such as unifying data types, dealing with spelling errors, disambiguating data, parsing to standard format. Data transformation is often the most complicated part. It is also the longest step in ETL development.**  **Data loading: The last step in ETL is to load the converted data near the target data warehouse.** |
| *Process 2* how to manage data warehouse |
| **Before developing an effective management strategy, the DTC must first investigate requirements. Between user-friendliness and data security, DTC can take a middle ground in managing its data.**  **Data warehouse jobs perform maintenance tasks, and the management package synchronization process imports management packages from Service Manager and defines how these management packages shape structure, move data, and replicate reports of data warehouse and reports. After these management packages are synchronized between the Service Manager and the data warehouse, the data is retrieved and reports are deployed for use by users.** |
| Prepare data warehouse user interface (You must prepare data table for each stakeholder, identify primary and foreign keys, link the table, identify the relationship and identify dimension and fact tables) |
| *Prepare warehouse data security strategy* |
| •       Identify who is accountable for security  **Safety should be the primary concern of all stakeholders (students, faculty, staff, visitors). Without a security and security policy, database users will be reluctant to enter their personal details. DTC management should encourage staff, faculty, students and visitors to take responsibility for protecting database operations, thereby minimizing potential security risks.** |
| •       How to test the warehouse to identify security issues  **1. Analyze source files: where the data comes from, how it is manipulated, and where it is stored.**  **2. Develop strategies and test plans: Incremental testing is the best way to test the data warehouse, with segmented tables of data capture followed by incremental tables, basic history tables, BI views, and so on.**  **3, test development and execution, test classification and test schedule can ensure that they also have a clear concept of test status.**  **Use web applications and database firewalls to identify security issues, block all pornographic websites and harmful content websites for students and staff, and block all websites without valid digital certificates.**  **Identify security issues by encrypting data.** |
| •       How to manage the security patches  **Update the site regularly and download from a trusted official website if necessary. You can also choose a patch management tool that fits your requirements.** |
| •       Where to store all the security issues  **Security questions should be kept in the information system for easy retrieval and updating.** |
| **Conclusion**  **My security strategy is to determine security accountability to stakeholders, then test the data warehouse, patch it, and locate security issues.** |