- Assignment 2 introduction
- Assignment 1 feedback
- Guest Speakers
  - Adam Li
  - Phil Watt

ISYS90086 Data Warehousing

Overhill Winery Case Study

Assignment 2

Overhill Winery

- Medium-sized boutique winery
- Goals
  - Managing future growth
  - Understanding customers, products, markets, and sales agents
- Operations
  - Growing grapes
  - Producing wine
  - Selling wine

- You have designed and built the data warehouse for Overhill Winery

- Characteristics of your design…

  - Fact Tables

    - Sales

      - Grain = Transaction (1 row per item sold)…

        » Why – because we are analysing actual sales

          » A customer can buy a particular product multiple times per day

    - Agent Commission

      - Grain = Monthly

        » As we are analyzing things like

          » Total earnings, number of sales, total amount sold etc.

- Characteristics of your design…
  - Conformed Dimensions
    - Product
    - Date
    - Customer
    - Market
    - Sales Agent

  - So when do you see an Exemplar design
    - After the 2nd assignment is handed in – as we are asking you to reflect on your design as a consequence of doing the ETL.

- ## Now What?
  - We need to do the loading of the data from the transactional systems into the data warehouse
    - We have done the extract on your behalf
      - You will receive in the zip file
        » Production System - The product and ProductionHistory data as .xlsx files
        » Sales System – 3 snapshots of the Customers over time, plus the Sale, SaleItem, Product, Market and Sales Agent data as .xlsx files
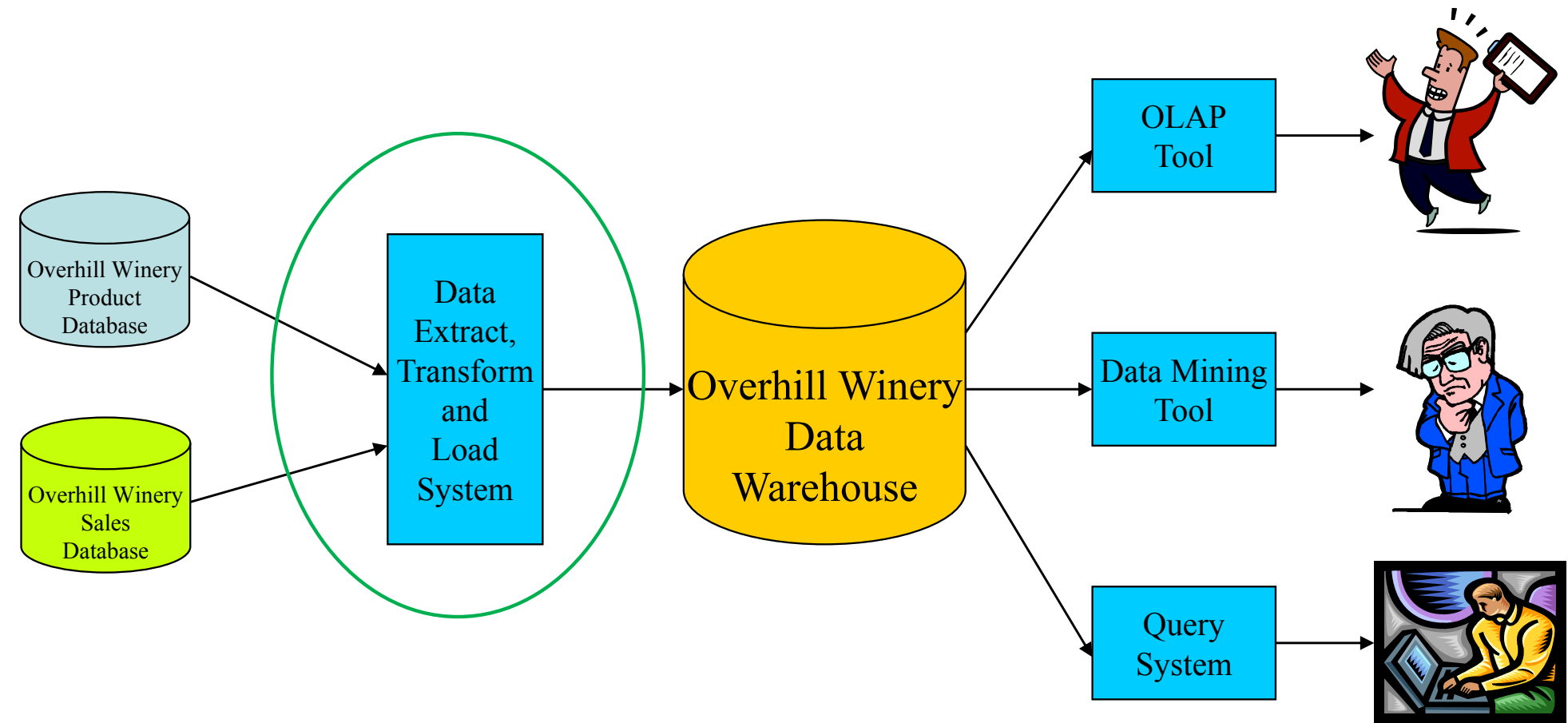   » The date dimension as a xlsx file

- **Your Job!**
  - Get the data into a format appropriate for insertion into the Data warehouse that you designed.

  - Remember the DW must support the 5 areas of concern for the business
  - Who are the key customers?
  - Which products are the most profitable?
  - Which market is the most profitable?
  - Which time periods are the most profitable?
  - Who are the key sales agent?

- You might realise your design needs improvement
  - So you might need to redesign

- ETL is hard and time consuming
  - Manage your time carefully – it's a 25 mark assignment
    - 80% of the stuff gets done in 20% of the time
      - Up to you as to the time you use

- Remember you are reporting on the process and the issues that you have in the ETL process
  - You have to submit a report AND
  - You have to submit a zip file of the transformations

- Report that describes and justifies the ETL process
- Around 2000 words in length
- Should include
  - Executive summary
  - Design of the ETL Process
    - Discuss and justify how you went about it for each fact and dimension table
  - Design of the data warehouse
    - If you did some redesign work on the DW what did you have to do and why
  - Data Dictionary
    - The TYPES of ADDITIONAL metadata you need to store as a result of the ETL process – need to describe (not on a field by field basis)
    - Don't repeat the DD from ass 1 – this is different, we are not asking you to produce a full DD now
  - Appendix 1: Work Breakdown (HURDLE)