TRƯỜNG ĐẠI HỌC KINH TẾ

**KHOA THỐNG KÊ – TIN HỌC**

–––––––––––––––––––––––––––––––

**BÁO CÁO THỰC TẬP NGHỀ NGHIỆP**

**TÊN ĐỀ TÀI**

**TRỰC QUAN HÓA DỮ LIỆU TRÊN DASHBOARD CHO KHÁCH HÀNG**

**Sinh viên thực hiện:** Nguyễn Hữu Hiếu Thiên Lớp : 44K21.2

**Giảng viên hướng dẫn:** Cao Thị Nhâm

**Đà Nẵng, 18/07/2021**

**DANH SÁCH THÀNH VIÊN THAM GIA**

|  |  |  |  |
| --- | --- | --- | --- |
| **TT** | **Họ và tên** | **Lớp** | **Nhiệm vụ** |
| 1 | Nguyễn Hữu Hiếu Thiên | 44K21.2 | Trưởng nhóm |

# LỜI CẢM ƠN

Để hoàn thành chuyên đề báo cáo thực tập này trước hết em xin gửi đến quý thầy cô giáo trong khoa Thống kê – Tin học trường Đại học Kinh tế Đại học Đà Nẵng lời cảm ơn chân thành.

Em xin chân thành cảm ơn Ban Lãnh Đạo, đặc biệt là anh Trần Thiên Quốc Tổng đã tạo điều kiện thuận lợi cho em được tìm hiểu thực tiễn trong suốt quá trình thực tập tại công ty.

Đặc biệt, em xin gửi đến cô Cao Thị Nhâm, người đã tận tình hướng dẫn, giúp đỡ em hoàn thành chuyên đề báo cáo thực tập này lời cảm ơn sâu sắc nhất.

Cuối cùng em xin cảm ơn các anh chị trong Công ty Cổ phần Công nghệ Dữ liệu DSA đã giúp đỡ, cung cấp những số liệu thực tế để em hoàn thành tốt chuyên đề thực tập nghề nghiệp này.

Vì kiến thức bản thân còn hạn chế nên trong quá trình thực tập này em không tránh khỏi những sai sót, kính mong nhận được những ý kiến đóng góp từ cô cũng như quý công ty.

# LỜI CAM ĐOAN

Em xin cam đoan dự Án **Trực Quan Hóa Dữ Liệu Trên Dashboard Cho Khách Hàng** là kết quả nghiên cứu độc lập dưới sự hướng dẫn của giáo viên Cao Thị Nhâm và mentor Trần Thiên Quốc Tổng. Ngoài ra không có bất cứ sự sao chép nào của người khác.

Dự án, nội dung báo cáo là sản phẩm mà em đã nỗ lực nghiên cứu trong quá trình học tập cũng như tìm hiểu trên ở nhiều phương tiện. Các dữ liệu, kết quả trình bày trong báo cáo là hoàn toàn trung thực. Em xin cam kết đề cương chưa từng được công bố trước đây. Em xin chịu hoàn toàn trách nhiệm, kỷ luật của Bộ môn và Nhà trường đề ra nếu có vi phạm.

# MỤC LỤC

[LỜI CẢM ƠN i](#_Toc77522897)

[LỜI CAM ĐOAN ii](#_Toc77522898)

[MỤC LỤC iii](#_Toc77522899)

[LỜI MỞ ĐẦU 1](#_Toc77522900)

[Chương 1 CƠ SỞ LÝ THUYẾT 3](#_Toc77522901)

[1.1 Tổng quan về BI 3](#_Toc77522902)

[1.1.1. Khái niệm 3](#_Toc77522903)

[1.1.2. Các thành phần 3](#_Toc77522904)

[1.1.3. Các chức năng của BI 5](#_Toc77522905)

[1.1.4. Các công nghệ hỗ trợ BI 5](#_Toc77522906)

[1.2 Tổng quan về Python 6](#_Toc77522907)

[1.2.1 Khái niệm 6](#_Toc77522908)

[1.2.2 Các tính chất 6](#_Toc77522909)

[1.2.3 Tính năng của Python 7](#_Toc77522910)

[1.3 Tổng quan về Microsoft SQL 8](#_Toc77522911)

[1.3.1 Khái niệm 8](#_Toc77522912)

[1.3.2 Các thành phần cơ bản trong SQL server 9](#_Toc77522913)

[Chương 2 Tổng quan về công nghệ 11](#_Toc77522914)

[và công cụ sử dụng 11](#_Toc77522915)

[2.1 Tổng quan về công nghệ 11](#_Toc77522916)

[2.2 Công cụ sử dụng 15](#_Toc77522917)

[Chương 3 PHân tích dữ liệu 15](#_Toc77522918)

[3.3.1 Hiểu nghiệp vụ khách hàng 17](#_Toc77522919)

[3.3.2 Hiểu yêu cầu khách hàng 17](#_Toc77522920)

[3.3.3 Hiểu dữ liêu khách hàng 17](#_Toc77522921)

[3.3.4 Thiết kế Dashboard 18](#_Toc77522922)

[3.3.5 Viết thủ tục 18](#_Toc77522923)

[3.3.6 Xây dựng Dashboard 19](#_Toc77522924)

[Chương 4 KẾT QUẢ 20](#_Toc77522925)

[KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 21](#_Toc77522926)

[Tài Liệu Tham Khảo 22](#_Toc77522927)

[PHỤ LỤC (nếu có) 23](#_Toc77522928)

DANH MỤC HÌNH ẢNH

[Hình 1. 1 Ngôn ngữ lập trình Python 8](#_Toc77519278)

[Hình 1. 2 Microsoft SQL Server 10](file:///C:\Users\Admin\Desktop\BaoCao_TNN_test.docx#_Toc77519279)

[Hình 2. 1 Thư viện DASH trong python 12](file:///C:\Users\Admin\Desktop\BaoCao_TNN_test.docx#_Toc77519283)

[Hình 2. 2 Thư viện Pandas trong python 13](file:///C:\Users\Admin\Desktop\BaoCao_TNN_test.docx#_Toc77519284)

[Hình 2. 3 Thư viện plotly trong python 14](file:///C:\Users\Admin\Desktop\BaoCao_TNN_test.docx#_Toc77519285)

[Hình 3. 1 Bản thiết kế Dashboard 18](#_Toc77519370)

DANH MỤC BẢNG BIỂU

[Bảng 4.1 Kiến trúc 5](#_Toc74235471)

DANH MỤC CÁC TỪ VIẾT TẮT

**BI:** Business Intelligence

**DW:** Data warehouse

**CSDL**: Cơ sở dữ liệu

# LỜI MỞ ĐẦU

1. **Mục tiêu của đề tài**

* Nghiên cứu ngôn ngữ lập trình Python(DASH)
* Nghiên cứu hệ quản trị CSDL Microsoft SQL Server 2019
* Xây dựng Dashboard, trực quan hóa dữ liệu
* Tương tác với dữ liệu được trực quan hóa

1. **Nhiệm vụ của đề tài**

* Nghiên cứu ngôn ngữ lập trình python (Thư viện Dash)
* Nghiên cứu Microsoft SQL Server 2019
* Phân tích bài toán quản lý dữ liệu:

*Công nợ phải thu***:**

* Tổng công nợ phải trả đến thời điểm hiện tại
* Top 30 NCC có công nợ lớn nhất.
* Danh sách công nợ quá hạn.
* Công nợ theo hợp đồng.
* Công nợ theo cơ cấu: hàng hóa, dịch vụ, khác.
* Tình hình thanh toán theo nhà cung cấp: đã mua bao nhiêu, thanh toán bao nhiêu, còn lại bao nhiêu.
* Thiết kế Dashboard và thực hiện vẽ dashboard bằng ngôn ngữ lập trình python
* Xây dựng chức năng tương tác 2 chiều với các biểu đồ trong dashboard

1. **Kết cấu của đề tài**

Đề tài được tổ chức gồm 4 chương nội dung và phần kết luận.

* **Chương 1**: Cơ sở lý thuyết
* **Chương 2**: Tổng quan về công nghệ và công cụ sử dụng
* **Chương 3**: Phân tích dữ liệu
* **Chương 4**: Kết quả
* Kết luận và hướng phát triển

# CƠ SỞ LÝ THUYẾT

## Tổng quan về BI

### Khái niệm

* BI là một qui trình có tích hợp công nghệ mà các doanh nghiệp dùng để kiểm soát khối lượng dữ liệu khổng lồ đến từ nhiều nguồn khác nhau và khai thác nguồn dữ liệu đó giúp cho họ có thể đưa các các quyết định hiệu quả hơn trong hoạt động kinh doanh của mình.

### Các thành phần

* BI chủ yếu gồm các thành phần chính như sau :
* *Data Sources:*
* Là cơ sở dữ liệu thô (thường là cơ sở dữ liệu quan hệ) đến từ nhiều nguồn khác nhau như các ứng dụng business như Human Resource Management (HRM), Customer relationship management (CRM), phần mềm bán hàng, website thương mại điện tử…
* Có thể là bất cứ hệ quản trị cơ sở dữ liệu nào như MySQL, Oracle, MSSQL, DB2, …
* Thường được thiết kế theo mô hình cơ sở dữ liệu quan hệ ( vì dạng mô hình này đang rất phổ biến trong thực tế )
* *Data Warehouse:*
* Là cơ sở dữ liệu được thiết kế theo mô hình khác với CSDL quan hệ và là nơi lưu trữ dữ liệu lâu dài của tổ chức
* Dữ liệu của DW chỉ có thể đọc, ko ghi hay update được và chỉ được update bởi gói ETL chuyển đổi dữ liệu từ Data Sources vào Data Warehouse.
* *Integrating Server:*

Chịu trách nhiệm trung gian vận hành gói ETL để chuyển đổi dữ liệu từ Data Sources vào Data Warehouse.

* *Analysis Server*:
* Chịu trách nhiệm thực thi các Cube được thiết kế dựa trên các Dimension dữ liệu và tri thức nghiệp vụ
* Cube chịu trách nhiệm nhận input data từ DW và thực thi theo nghiệp vụ định nghĩa sẵn để trả về output.
* *Reporting Server:*
* Thực thi các report với output nhận được từ Analysis Server.
* Nơi quản trị tập trung các report trên nền web, các report này có thể được attach vào ứng dụng web, hay application
* *Data Mining*
* Là quá trình trích xuất thông tin dữ liệu đã qua xử lý (phù hợp với yêu cầu riêng của doanh nghiệp) từ Data Warehouse rồi kết hợp với các thuật toán để đưa ra ( hoặc dự đoán ) các quyết định có lợi cho việc kinh doanh của doanh nghiệp.
* Đây là một quá trình quan trọng trong BI, thông thường một doanh nghiệp muốn sử dụng giái pháp BI thường kèm theo về Data Mining.
* *Data Presentation*
* Tạo ra các báo cáo, biểu đồ từ quá trình data mining để phục vụ cho nhu cầu của người dùng cuối.

### Các chức năng của BI

* Xử lý phân tích trực tuyến (Online analytical processing)
* Khai thác dự liệu (Data mining)
* Hỗ trợ quyết định (Decision support)
* Truy vấn và báo cáo (Query and reporting)
* Phân tích thống kê (Statistical analysis)
* Phân tích dự đoán và phân tích theo quy định (Forecasting)

### Các công nghệ hỗ trợ BI

* Kho dữ liệu (Data warehousing)
* Hệ thống hỗ trợ ra quyết định (Decision support systems)
* Công cụ khai thác và phân tích dữ liệu (Data mining and analytics tools)
* Hệ thống hoạch định nguồn lực doanh nghiệp (Enterprise resource Planning (ERP) systems)
* Công nghệ truy vấn và lập báo cáo (Query and report writing technologies)
* Quản lý quan hệ khách hàng (Customer relation management)

## Tổng quan về Python

### Khái niệm

* Python là ngôn ngữ lập trình hướng đối tượng thông dụng dùng để viết các tiện ích hệ thống. Nó cũng được sử dụng như ngôn ngữ kết dính đóng vai trò tích hợp C và C++.
* Python là ngôn ngữ có hình thức khá đơn giản và rõ ràng, do đó tạo nên sự dễ dàng tiếp cânh cho những lập trình viên mới bắt đầu.
* Ban đầu, Python được phát triển để chạy trên nền Unix, nhưng rồi theo thời gian, nó đã được mở rộng sang mọi hệ điều hành từ MS-DOS đến Mac OS, OS/2, Windows, Linux và các hệ điều hành khác thuộc họ Unix.

### Các tính chất

* **Interpreted**: nhờ chức năng thông dịch mà trình thông dịch (Interpreter) của Python có thể xử lý lệnh tại thời điểm chạy chương trình (runtime). Nhờ đó mà không cần biên dịch chương trình trước khi thực hiện nó (tương tự như Perl và PHP).
* **Interactive**: tính năng tương tác của Python giúp tương tác trực tiếp với trình thông dịch của nó ngay tại dấu nhắc lệnh. Cụ thể: Có thể thực hiện lệnh một cách trực tiếp tại dấu nhắc của Python.
* **Object-Oriented**: Python hỗ trợ mạnh cho phong cách lập trình hướng đối tượng và kỹ thuật lập trình gói mã trong đối tượng.
* **Beginner's Language**: Python được xem là ngôn ngữ lập trình dành cho những ai mới làm quen với việc lập trình trên máy tính, nhưng nó hỗ trợ mạnh cho việc phát triển nhiều loại ứng dụng khác nhau, từ các chương trình xử lý văn bản đơn giản đến các ứng dụng web, đến các chương trình game, …

### Tính năng của Python

* Python là một ngôn ngữ lập trình dễ học, dễ đọc. Python tăng cường sử dụng từ khóa tiếng Anh, hạn chế các ký hiệu và cấu trúc cú pháp so với các ngôn ngữ khác.
* Mã nguồn của Python tương đối dễ để bảo trì và duy trì và có khả năng mở rộng.
* Python có một tiêu chuẩn thư viện rộng, nền tảng tương thích trên nhiều hệ điều hành hiện nay như: UNIX, Windows, và Macintosh. Đây cũng là một trong những điểm mạnh đáng kể của Python.
* Python có thể được sử dụng như ngôn ngữ script, hoặc ngôn ngữ biên dịch, nhờ đó có thể build các chương trình lớn trên nó.
* Khi chế độ Interactive, có thể nhập kết quả từ các đầu cuối khác nhau vào chương trình Python, do đó mà việc test hay debug lỗi code trở nên đơn giản hơn.
* Python cho phép người dùng tích hợp vào các module để có thể sử dụng trong các chương trình khác. Nó cũng cung cấp sẵn một tập hợp các modules chuẩn mà lập trình viên có thể sử dụng lại trong chương trình của họ. Các module này cung cấp nhiều chức năng hữu ích, như các hàm truy xuất tập tin, các lời gọi hệ thống, hỗ trợ lập trình mạng (socket),…
* Python cung cấp giao diện cho tất cả các cơ sở dữ liệu thương mại lớn.
* Có thế dễ dàng tích hợp với C, C++, COM, CORBA, ActiveX, Java.



Hình 1. Ngôn ngữ lập trình Python

## Tổng quan về Microsoft SQL

### Khái niệm

* SQL Server chính là một hệ quản trị dữ liệu quan hệ sử dụng câu lệnh SQL để trao đổi dữ liệu giữa máy cài SQL Server và máy Client. Một Relational Database Management System – RDBMS gồm có: databases, datase engine và các chương trình ứng dụng dùng để quản lý các bộ phận trong RDBMS và những dữ liệu khác.
* SQL server chính là cụm từ viết tắt của Structure Query Language được sử dụng nhiều trong các lĩnh vực bởi chức năng quản lý dữ liệu. Các ngôn ngữ cấp cao như: [Visual C](https://www.visualstudio.com/vs/cplusplus/), [Oracle](https://www.oracle.com/database/index.html), Visual Basic,… đều có trình hỗ trợ là SQL. Những ứng dụng khi chạy phải sử dụng SQL khi người dùng truy cập tới cơ sở dữ liệu thì không cần sử dụng trực tiếp SQL.

### Các thành phần cơ bản trong SQL server

* **Database Engine**: Đây là một engine có khả năng chứa dữ liệu ở các quy mô dưới dạng support và table. Ngoài ra, nó còn có khả năng tự điều chỉnh ví dụ: trả lại tài nguyên cho ệ điều hành khi một user log off và sử dụng thêm các tài nguyên của máy khi cần.
* **Integration Services**: là tập hợp các đối tượng lập trình và các công cụ đồ họa cho việc sao chép, di chuyển và chuyển đổi dữ liệu.  Khi bạn làm việc trong một công ty lớn thì dữ liệu được lưu trữ ở nhiều nơi khác nhau như được chứa trong: Oracle, SQL Server, DB2, Microsoft Access,… và bạn chắc chắn sẽ có nhu cầu di chuyển dữ liệu giữa các server này. Ngoài ra, bạn còn muốn định dạng dữ liệu trước khi lưu vào database. Chắc chắn Integration Services sẽ giúp bạn giải quyết được công việc này dễ dàng.
* **Analysis Services**: Đây là một dịch vụ phân tích dữ liệu rất hay của Microsoft. Dữ liệu khi được lưu trữ vào trong database mà bạn không thể lấy được những thông tin bổ ích thì coi như không có ý nghĩa gì. Chính vì thế, công cụ này ra đời giúp bạn trong việc phân tích dữ liệu một cách hiệu quả và dễ dàng bằng cách dùng kỹ thuật khai thác dữ liệu – datamining và khái niệm hình khối nhiều chiều – multi dimendion cubes.
* **Notification Services**: Dịch vụ thông báo này là nền tảng cho sự phát triển và triển khai các ứng dụng soạn và gửi thông báo. Ngoài ra, dịch vụ này còn có chức năng gửi thông báo theo dịch thời đến hàng ngàn người dăng ký sử dụng trên nhiều loại thiết bị khác nhau.
* **Reporting  Services**: là một công cụ tạo, quản lý và triển khai báo cáo bao gồm: server và client. Ngoài ra, nó còn là nền tảng cho việc phát triển và xây dựng các ứng dụng báo cáo.
* **Full Text Search Service**: là một thành phần đặc biệt trong việc truy vấn và đánh chỉ mục dữ liệu văn bản không cấu trúc được lưu trữ trong các cơ sở dữ liệu SQL Server.
* **Service Broker**: là một môi trường lập trình cho việc tạo ra các ứng dụng trong việc nhảy qua các Instance.



Hình 1. Microsoft SQL Server

# Tổng quan về công nghệ

# và công cụ sử dụng

## Tổng quan về công nghệ

* 1. **Thư viện DASH**
* Dash của Python được viết trên Flask, Plotly.js và React.js, Dash lý tưởng để xây dựng các ứng dụng trực quan hóa dữ liệu với giao diện người dùng tùy chỉnh cao bằng Python thuần túy, phù hợp làm việc với dữ liệu bằng Python.
* Dash là mã nguồn mở và các ứng dụng của nó chạy trên trình duyệt web.
* Mục tiêu là tạo trang tổng quan ứng dụng web với các kỹ năng phát triển web tối thiểu hoặc không có như Javascript. Mặc dù Dash Python cho phép tùy chỉnh tốt hơn với việc sử dụng ngăn xếp Javascript, HTML, CSS
* Dash đi kèm với thư viện các thành phần cốt lõi, cho phép dễ dàng khai báo các thành phần HTML phổ biến như Dropdowns hoặc Graphs.
* Một điều thuận tiện khi Plotly graph\_objects được kết nối với Dash là bạn tự động nhận được các tính năng của Plotly mà không cần thêm mã. Điều này bao gồm thu phóng, tính năng di chuột, so sánh di chuột, tùy chọn tải xuống, chụp ảnh,…

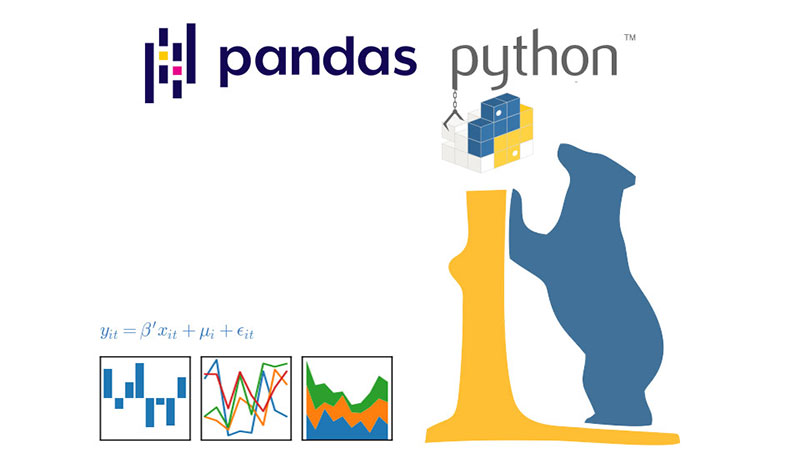
*Bố cục ứng dụng dash*

* Một ứng dụng Dash thường bao gồm hai phần. Phần đầu tiên là bố cục và mô tả ứng dụng và phần thứ hai là mô tả khả năng tương tác của ứng dụng.
* ***dash\_html\_components:*** cung cấp các lớp HTML cho phép tạo nội dung HTML bằng Python. Ngoài ra, có thể tạo các thành phần tùy chỉnh của riêng mình bằng cách sử dụng Javascript và React Js.
* ***dash\_core\_components:*** tạo một danh sách thả xuống, chuyển các tùy chọn dưới dạng danh sách các từ điển, các nút radio sử dụng RadioItems, gọi Checklist tạo các hộp kiểm tra và Input – văn bản đầu vào trong ứng dụng.
* Để có thể tương tác trong ứng dụng Dash, chúng ta cần nhập Input và Output từ dash.dependencies. Để kích hoạt, Dash này cung cấp một trình trang trí @app để có thể liên kết một hạn gọi lại id đã tạo và trường đầu vào HTML.



Hình 2. Thư viện DASH trong python

* 1. **Thư viện PANDAS**
* Thư viện pandas trong Python là một thư viện mã nguồn mở, hỗ trợ đắc lực trong thao tác dữ liệu. Đây cũng là một công cụ phân tích và xử lý dữ liệu mạnh mẽ của ngôn ngữ lập trình Python.
* Thư viện này sử dụng một cấu trúc dữ liệu riêng là DataFrame. Đồng thời, pandas cũng cung cấp rất nhiều chức năng xử lý và làm việc trên cấu trúc dữ liệu này.
* cho phép đọc/ ghi dữ liệu giữa bộ nhớ và nhiều định dạng file: csv, text, excel, sql,…
* Liên kết dữ liệu thông minh, xử lý được trường hợp dữ liệu bị thiếu. Tự động đưa dữ liệu lộn xộn về dạng có cấu trúc.
* Dễ dàng thay đổi bố cục của dữ liệu.
* Tích hợp cơ chế trượt, lập chỉ mục, lấy ra tập con từ tập dữ liệu lớn.
* Có thể thêm, xóa các dữ liệu.
* Tập hợp hoặc thay đổi dữ liệu với group by cho phép thực hiện các toán tử trên tập dữ liệu.
* Hiệu quả cao trong kết hợp các tập dữ liệu.



Hình 2. Thư viện Pandas trong python

* 1. **Thư viện PLOTLY**
* Plotly là một thư viện hướng mã nguồn mở, hỗ trợ người dùng trực quan hóa dữ liệu tương tác. Đồng thời, thư viện này cho phép vẽ đồ thị tương tác không những trực quan hơn mà còn thuyết phục người nghe hơn.
* Hỗ trợ hơn 40 loại bảng biểu khác nhau cho rất nhiều ứng dụng về thống kê, tài chính, địa lý, khoa học và các ứng dụng đời thực khác.
* Cho phép người dùng Python tạo ra các thiết kế web trực quan, sinh động để biểu thị trên nhiều loại cửa sổ như Visual Studio Code, Jupyter.
* Lưu trữ trong tập HTML độc lập hoặc tích hợp trong các ứng dụng web thuần Python thông qua Dash.

Hình 2. Thư viện plotly trong python

## Công cụ sử dụng

* Công cụ phổ biến dùng để Viết plotly Dash : Visua studio code, Sublime text3, Pycharm, Visua studio…
* Công cụ xử lý dữ liệu : SQL SERVER, mysql…

# PHân tích dữ liệu

1. **Mô tả bài Toán**

* Những khó khăn của doanh nghiệp trong việc quản lý nguồn dữ liệu lớn, là sự trộn lẫn các dữ liệu không liên quan với nhau. Việc này có thể làm doanh nghiệp mất thời gian trong việc sắp xếp các dữ liệu, dẫn đến việc mất các dữ liệu quan trọng. Nhưng khi các dữ liệu được trình bày trên một Dashboard, thì những dữ liệu cần thiết sẽ liên quan mật thiết, chặt chẽ với nhau. Khách hàng dễ dàng phân tích dữ liệu cụ thể một cách chuyên sâu hơn. Dashboard ngoài việc dễ hiểu nhờ sự sắp xếp hợp lý, thì còn giúp người dùng dễ dàng tập trung nhiều hơn vào những chỉ số liên quan đến các chiến dịch đề ra.
* Xuất phát từ những vấn đề trên, nhóm được giao để thực hiện “Trực quan hóa dữ liệu trên Dashboard cho khách hàng”.

1. **Thiết kế cơ sở dữ liệu**

* **Bảng dmhd**:
* *Ma\_hd*: mã hóa đơn
* *Ten\_hd*: tên hóa đơn
* *Ma\_kh*: Mã khách hàng
* *Ngay\_hd*: ngày hóa đơn
* *So\_hd*: số hóa đơn
* **Bảng dmkh**:
* *Ma\_kh:* Mã khách hàng
* *Ten\_kh:* Tên khách hàng
* *Dia\_chi:* Địa chỉ
* *Ngan\_hang:* Ngân hàng
* **Bảng dmct**:
* *Ma\_phan\_he:* Mã phân hệ
* *Ma\_ct:* Mã chứng từ
* *Ten\_ct:* Tên chứng từ
* *So\_ct:* số chứng từ
* **Bảng cttt20:**
* *Ma\_ct:* Mã chứng từ
* *So\_ct:* số chứng từ
* *Ma\_kh:* Mã khách hàng
* *Ma\_bp:*Mã bộ phận
* *Ngay\_ct:* ngày chứng từ
* *T\_tien2:* tiền phải thu
* *T\_thue:* thuế

1. **Các Bước phân tích dữ liệu**

### Hiểu nghiệp vụ khách hàng

Tìm hiểu nghiệp vụ thông qua phần mềm ERP của khách hàng.

Daily meeting trình bày những thông tin đã tìm hiểu được.

Meeting online với khách hàng để trao đổi về nghiệp vụ đã hiểu đúng hay chưa?

### Hiểu yêu cầu khách hàng

Sau khi có được yêu cầu từ khách hàng, cần phải meeting với khách hàng để xác định lại là đã hiểu đúng yêu cầu của khách hàng chưa , nếu đúng thì mới có thể làm tiếp những bước sau

### Hiểu dữ liêu khách hàng

Tìm hiểu dữ liệu mà khách hàng cung cấp để biết những dữ liệu nào có liên quan đến nghiệp vụ mà khách hàng yêu cầu. sau đó meeting với khách hàng để xác định đã tìm đúng dữ liệu chưa.

### Thiết kế Dashboard

Thiết kế 1 bản dashboard mẫu để sau đó tiến hành làm dashboard theo mẫu



Hình 3. Bản thiết kế Dashboard

### Viết thủ tục

* Viết các thủ tục cần thiết để sử dụng khi xây dựng Dashboard
* Tổng doanh thu
* Tổng tiền đã thu
* Còn lại phải thu
* Doanh thu và tiền theo thời gian
* Công nợ phải thu theo thời gian
* Top 30 khách hàng có công nợ lớn nhất
* Tình hình công nợ phải thu theo từng khách hàng
* Cơ cấu doanh thu theo dịch vụ
* Nợ phải thu theo chứng từ
* Nợ phải thu theo hợp đồng
* Lấy giá trị cho các dropdown list
* Gộp tất cả các thủ tục lại thành 1 thủ tục chính

### Xây dựng Dashboard

Sau khi đã hoàn thành tất cả các bước ở trên thì bắt đầu tiền hành làm layout cho dashboard bằng các thư viện dash,pandas, plotly trong python và sau đó code các chứng năng tương tác cho dashboard.

* Kết nối dữ liệu từ SQL Server sang Python bằng Pymssql
* Tạo các hàm để sử dụng
* Tạo Layouts cho các biểu đồ
* Tạo tương tác cho Dashboard
* Tạo CSS để chỉnh sửa giao diện
* Import tất cả các file vào file chính.

# KẾT QUẢ

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

# Tài Liệu Tham Khảo

Basic Dash Callbacks. *Plotly.* [Online] https://dash.plotly.com/basic-callbacks.

Dash DataTable. *Plotly.* [Online] https://dash.plotly.com/datatable.

Dash Layout. *Plotly.* [Online] https://dash.plotly.com/layout.

Plotly Python Open Source Graphing Library. *Plotly.* [Online] https://plotly.com/python/.

# PHỤ LỤC (nếu có)