A Prospective Data Warehouse for a Blood Donation Clinic

Daniel G. Barstow

Department of Computer Science and Information Systems

Elmhurst University

Dr. Dean Jensen

9 December 2022

Table of Contents

Section 1 - Requirements

Business Description

Blood donation clinics acquire blood through donors to give to hospitals for car accident victims, organ transplant patients, and people with cancer. Donors must first register and complete a physical exam before donating to make sure that they don't have any blood borne illnesses such as HIV or hepatitis. Donors list their health history, any medications they take, and any sexual activity which may affect their eligibility to donate. After the donation, the clinic performs a series of tests on the blood to make sure it's suitable for transfusion. If a positive result comes back on any of the tests, the blood is discarded. If the blood passes all the tests, the clinic processes the blood, separating red blood cells, platelets, and plasma. Each of these must be refrigerated, and each has a limited shelf-life: 42 days for red blood cells, 5 days for platelets, and 1 year for plasma. When a request for a particular type of blood is made by a hospital, it is shipped off to be used in a transfusion. Since the clinic rarely has many donors, blood sits on shelves more often than used for transfusions, so a data warehouse has been proposed as a solution for keeping better track of the blood they collect.

Business Requirements

- A few small blood donation clinics are trying to find a way to keep better track of the donations they receive to speed up the process of testing and shipping blood to local hospitals and to reduce the time donated blood spends on shelves.
- Donors must register and take a physical exam to be able to donate. If they have blood borne illnesses, engage in certain sexual activities, or take particular medications then they cannot donate.
- If a donor passes the examination, then they donate approximately 1 pint of blood (16 fl. oz.).
- Generally speaking, a donor can only donate once every 8 weeks. This isn't necessarily the case depending upon what they are donating for (red blood cells, platelets, plasma), but it suffices for this project.
- Since people have different blood types, they can only donate to someone who has the same blood type, with some exceptions, so the clinic must note the blood type.
- The clinic separates each constituent of the donated blood: red blood cells, platelets, and plasma.
- The blood is then kept in an inventory and each constituent has a limited shelf-life before it degrades. The clinic must track this and discard blood past its expiry date.
- The business wants to know the blood inventory breakdown of each clinic. Knowing this will help them determine what type of blood they need more of versus what they already have on their shelves.
- The business also wants to know if they are over-collecting certain kinds of blood over others. Knowing this will help them cut down on unnecessary donations and reduce the cost of keeping more blood on-hand.

- In the end, business users want to see a list of recent donors along with their medical histories and medications, how much of each blood type they have in inventory along with their expiry dates, and how much blood of differing types is collected throughout the year.

List of Questions the Business Might Ask:

1) How much of each blood type is there in inventory?
2) What is the breakdown of intended uses for the blood in inventory?
3) Which date had the most donations?
4) Is there any blood past its expiry date in inventory?
5) What are the phone numbers and email addresses of people with a certain blood type?
6) Which clinic has the most of a certain blood type on-hand?
7) How many medications does a particular donor take?
8) What are the phone numbers of people who take a certain medication?
9) What is the name of the clinic that took the most recent donation?
10) What are all the names of clinics that a particular donor has donated at before?

Conceptual Data Model

The conceptual data model includes 3 entities: a donor, the clinic accepting the donation, and the donation itself. The donor entity has 9 attributes which include basic contact information, an identifier, and information that affects their eligibility to donate such as medical history, medications they take, their sexual activity, and their blood type. The donation entity has 6 attributes: an identifier, the date of the donation, the blood type drawn, the amount of blood donated, the intended use for the blood (red blood cells, platelets, plasma), and the expiration date. The clinic entity has 5 attributes consisting of an identifier and basic contact information. The rudimentary conceptual data model for this project is seen in fig. 1.

Figure 1 – The Initial Conceptual Data Model

Database Requirements

Requirements Enforced by the Database

- Each DONATION is *donated* by exactly one DONOR. Conversely, each DONOR may *donate* many times, but not all at once, as there is an 8-week waiting period between DONATIONS.
- Each DONATION is *collected* by exactly one CLINIC. Conversely, any CLINIC may *collect* any number of DONATIONS.

Requirements not Enforced by the Database

- The database will not prevent DONATION records from being added which reference a DONOR who is not eligible, since the phlebotomists will look at this prior to taking the DONATION.
- The database will not prevent DONATION records from being added if the DONOR donates before the required 8-week waiting period has expired, because the phlebotomists will have already taken the DONATION.

Section 2 – Data Models

Logical Data Model

The logical data model consists of 3 entities with several attributes. The ER diagram for this data model is shown in fig. 2. The entities and attributes are:

Donor – an entity consisting of information about the person who registers to donate blood. It has 9 attributes.

- Donor ID – a 10-digit unique identifier associated with a particular donor.
- Name – a string containing the full name of the donor.
- Address – a string containing the address of the donor.
- Postal Code – a 5-digit number representing the postal/zip code of the donor.
- Phone Number – a 12-character string consisting of the area code, prefix, and line number for a particular donor.
- Blood Type – a 2- or 3-character string representing the donor's blood type.
- Medications – a string consisting of medications the donor is currently taking.
- Medical History – a string consisting of previous or current illnesses the donor has or has had in the past.
- Sexual Activity – a Boolean indicating relevant sexual activity the donor has engaged in that may or may not prevent them from being eligible to donate.

Donation – an entity consisting of information about a particular donation. It has 6 attributes.

- Donation ID – a 10-digit unique identifier associated with a particular donation.
- Donation Date – the date on which the donation was given.
- Blood Type Donated – a 2- or 3-character string representing the blood type donated.
- Amount of Blood Donated – a 10-character string representing the number of fluid ounces of blood donated. Since most donations amount to roughly a pint (16 fl. oz.), this is sufficient for the database.
- Intended Use – a 5- to 7-character string representing the intended use case for the blood donation (red blood cells, platelets, or plasma).
- Expiration Date – the date at which the blood donation can no longer be used for transfusions.

Clinic – an entity consisting of information about the clinic which is taking donations. It has 5 attributes.

- Clinic ID – a 4-digit unique identifier associated with a particular blood donation clinic.
- Name – a string consisting of the name of the blood donation clinic.
- Address – a string containing the address associated with a particular blood donation clinic.
- Postal Code – a 5-digit number representing the postal/zip code for the blood donation clinic.
- Phone Number – a 12-character string consisting of the area code, prefix, and line number for a particular blood donation clinic.

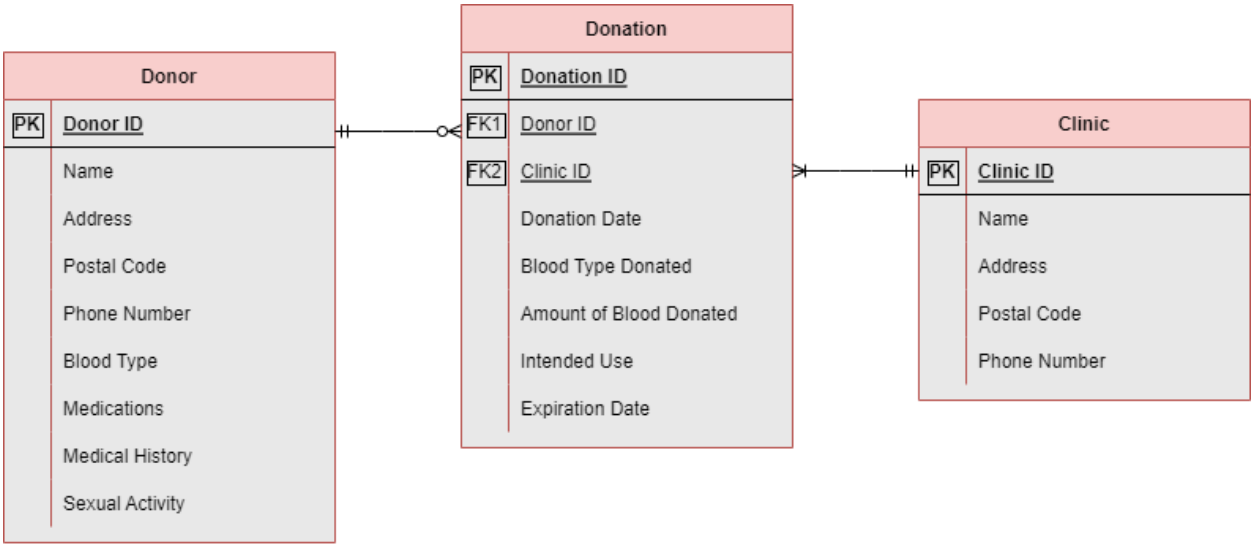

Figure 2 – The Initial Entity-Relationship Diagram

Section 3 – Physical Database

Un-normalized Tables

DONOR Table

Donor_ID	Name	Address	Postal_Code	Phone_Number	Blood_Type	Medications	Medical_History	Sexual_Activity

DONATION Table

Donation_ID	**Donor_ID**	**Clinic_ID**	Donation_Date	Blood_Type_Donated	Amount_of_Blood_Donated	Intended_Use	Expiration_Date

CLINIC Table

Clinic_ID	Name	Address	Postal_Code	Phone_Number

1st Normal Form

In 1st normal form, no two rows of data contain repeating groups of information. In other words, each set of columns must have a unique value.

- Since the Medications column in the DONOR table may contain multiple values, rather than listing the medications and dosages in a long piece of text, it was split off into its own table and the two were connected with a bridge table called DONOR_MEDICATION.
- The DONOR_MEDICATION table only contains the composite primary key Donor_Medication_ID, Medication_ID.
- The new MEDICATION table has a primary key, a name, a medication type, and a dosage.
- The resulting ER diagram is shown in fig. 3.

- While the Medical_History attribute being a string breaks this 1st normal form, it was decided to keep it as it is because this data warehouse is only designed for a small number of blood donation clinics.
- Also, an Email Address Attribute was added to the DONOR table.

New MEDICATION Table

Medication_ID	Medication_Name	Medication_Type	Dosage

Updated DONOR Table

Donor_ID	Donor_Medication_ID	Name	Address	Postal_Code	Phone_Number	Email_Address	Blood_Type	Medical_History	Sexual_Activity_Eligibility

NEW DONOR MEDICATION Table

Donor_Medication_ID	Medication_ID

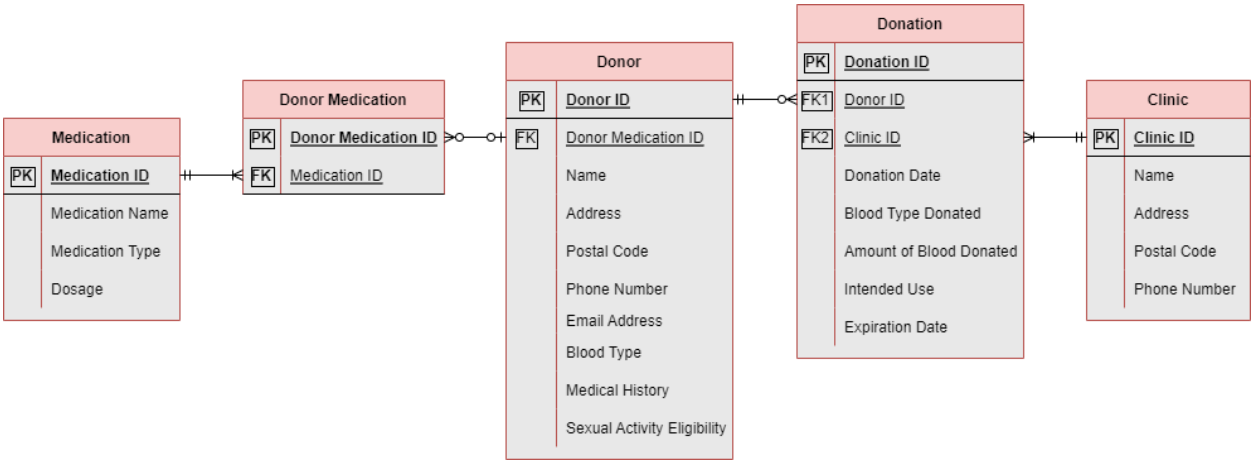

Figure 3 – The Normalized Entity-Relationship Diagram

2nd Normal Form

In 2nd normal form, there must not be any partial dependency of any column on the primary key. If there is a partial dependency, then the table must be split.

- No issues of partial dependency were found.

3rd Normal Form

In 3rd normal form, non-key columns directly depend on the primary key. In other words, there should be no case where a non-prime attribute is determined by another non-prime attribute.

- No issues of transitive dependency were found.

Normalized Tables with Imaginary Sample Data

DONOR

Donor_ ID	Donor_Medication_ID	Name	Address	Postal_ Code	Phone_Number	Email_Address	Blood_ Type	Medical_ History	Sexual_Activity _Eligibility
1111111111	0123456789	Bob Ferguson	89 W. Livermore St.	60175	1-265-651-8859	ferguson@yahoo.com	A+	n/a	1
2222222222	1234567890	Joan Sanders	1628 Northridge Ave.	60167	1-265-987-0647	sanders@gmail.com	O-	n/a	1
3333333333	2345678901	Emilio Sanchez	11 N. Water St.	60162	1-630-657-5984	sanchez@elmhurst.edu	AB-	Type-2 Diabetes	1
4444444444	3456789012	Yan Lee	25 E. Main St.	60177	1-630-549-8899	lee@outlook.com	B+	Smoker	1
5555555555	4567890123	Joe Brown	365W 872 Maple Dr.	60172	1-780-888-9049	brown@company.net	A-	Hepatitis	0

DONATION

Donation_ID	**Donor_ID**	**Clinic_ID**	Donation_Date	Blood_Type_Donated	Amount_of_Blood_Donated	Intended_Use	Expiration_Date
1212121212	1111111111	1111	11/1/22	A+	15	Platelets	11/6/22
2323232323	1111111111	1111	11/1/22	A+	16	RBC	12/12/22
3434343434	2222222222	2222	11/1/22	O-	14	RBC	12/12/22
4545454545	3333333333	3333	11/1/22	AB-	13	Plasma	11/1/23
5656565656	4444444444	3333	11/1/22	B+	15	Plasma	11/1/23

CLINIC

Clinic_ID	Name	Address	Postal_Code	Phone_Number
1111	St. Mary's Blood Clinic	49 W. Flurry Ave.	60174	1-630-499-1223
2222	East City Blood Drive	52 N. Elm Dr.	60119	1-630-159-5619
3333	Blood for Cash	529 Maple St.	60112	1-459-589-9992

DONOR MEDICATION

Donor_Medication_ID	**Medication_ID**
1111111111	1001001001
2222222222	2002002002
3333333333	3003003003

MEDICATION

Medication_ID	Medication_Name	Medication_Type	Dosage
1001001001	Lexapro	Antidepressant	10mg
2002002002	Adderall	Amphetamine	20mg
3003003003	Xanax	Benzodiazepine	0.5mg

SQL Code – Scripts to Create Database Tables in SQL Server

- During the creation of the tables, it was decided that the Amount_of_Blood_Donated attribute should be renamed to Amount_Donated_floz so that the column would be additive and wouldn't consist of any characters like "fl. oz.".
- Similarly, the Dosage attribute was changed to Dosage_mg to preserve additivity on the column as well.
- The schema diagram created from these tables in MS SQL Server is shown in fig. 4.

```
USE [MDS523_Barstow_Final_Project]
GO

/****** Object:  Table [dbo].[Clinic]    Script Date: 12/6/2022 11:01:23 AM ******/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Clinic](
	[Clinic_ID] [char](4) NOT NULL,
```

```
	[Name] [varchar](50) NOT NULL,
	[Address] [varchar](50) NOT NULL,
	[Postal_Code] [char](5) NOT NULL,
	[Phone_Number] [varchar](12) NOT NULL,
 CONSTRAINT [PK_Clinic] PRIMARY KEY CLUSTERED
(
	[Clinic_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
USE [MDS523_Barstow_Final_Project]
GO

/****** Object:  Table [dbo].[Donation]    Script Date: 12/6/2022 11:01:38 AM ******/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Donation](
	[Donation_ID] [char](10) NOT NULL,
	[Donor_ID] [char](10) NOT NULL,
	[Clinic_ID] [char](4) NOT NULL,
	[Donation_Date] [date] NOT NULL,
	[Blood_Type_Donated] [varchar](3) NOT NULL,
	[Amount_Donated_floz] [float] NOT NULL,
```

```
	[Intended_Use] [varchar](7) NOT NULL,
	[Expiration_Date] [date] NOT NULL,
 CONSTRAINT [PK_Donation] PRIMARY KEY CLUSTERED
(
	[Donation_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

ALTER TABLE [dbo].[Donation]  WITH CHECK ADD  CONSTRAINT [FK_Donation_Clinic] FOREIGN KEY([Clinic_ID])
REFERENCES [dbo].[Clinic] ([Clinic_ID])
GO

ALTER TABLE [dbo].[Donation] CHECK CONSTRAINT [FK_Donation_Clinic]
GO

ALTER TABLE [dbo].[Donation]  WITH CHECK ADD  CONSTRAINT [FK_Donation_Donor] FOREIGN KEY([Donor_ID])
REFERENCES [dbo].[Donor] ([Donor_ID])
GO

ALTER TABLE [dbo].[Donation] CHECK CONSTRAINT [FK_Donation_Donor]
GO
USE [MDS523_Barstow_Final_Project]
GO

/****** Object:  Table [dbo].[Donor]    Script Date: 12/6/2022 11:01:49 AM ******/
```

```
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Donor](
	[Donor_ID] [char](10) NOT NULL,
	[Donor_Medication_ID] [char](10) NULL,
	[Name] [varchar](50) NOT NULL,
	[Address] [varchar](50) NOT NULL,
	[Postal_Code] [char](5) NOT NULL,
	[Phone_Number] [varchar](12) NOT NULL,
	[Email_Address] [varchar](50) NOT NULL,
	[Blood_Type] [varchar](3) NOT NULL,
	[Medical_History] [varchar](5000) NULL,
	[Sexual_Activity_Eligibility] [bit] NOT NULL,
 CONSTRAINT [PK_Donor] PRIMARY KEY CLUSTERED
(
	[Donor_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO

ALTER TABLE [dbo].[Donor]  WITH CHECK ADD  CONSTRAINT [FK_Donor_Donor_Medication] FOREIGN
KEY([Donor_Medication_ID])
REFERENCES [dbo].[Donor_Medication] ([Donor_Medication_ID])
GO
```

```
ALTER TABLE [dbo].[Donor] CHECK CONSTRAINT [FK_Donor_Donor_Medication]
GO
USE [MDS523_Barstow_Final_Project]
GO

/****** Object:  Table [dbo].[Donor_Medication]    Script Date: 12/6/2022 11:01:58 AM ******/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Donor_Medication](
	[Donor_Medication_ID] [char](10) NOT NULL,
	[Medication_ID] [char](10) NOT NULL,
 CONSTRAINT [PK_Donor_Medication] PRIMARY KEY CLUSTERED
(
	[Donor_Medication_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO

ALTER TABLE [dbo].[Donor_Medication]  WITH CHECK ADD  CONSTRAINT
[FK_Donor_Medication_Medication] FOREIGN KEY([Medication_ID])
REFERENCES [dbo].[Medication] ([Medication_ID])
GO
```

```
ALTER TABLE [dbo].[Donor_Medication] CHECK CONSTRAINT [FK_Donor_Medication_Medication]
GO
USE [MDS523_Barstow_Final_Project]
GO

/****** Object:  Table [dbo].[Medication]    Script Date: 12/6/2022 11:02:08 AM ******/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Medication](
	[Medication_ID] [char](10) NOT NULL,
	[Medication_Name] [varchar](50) NOT NULL,
	[Medication_Type] [varchar](50) NULL,
	[Dosage_mg] [int] NOT NULL,
 CONSTRAINT [PK_Medication] PRIMARY KEY CLUSTERED
(
	[Medication_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
```

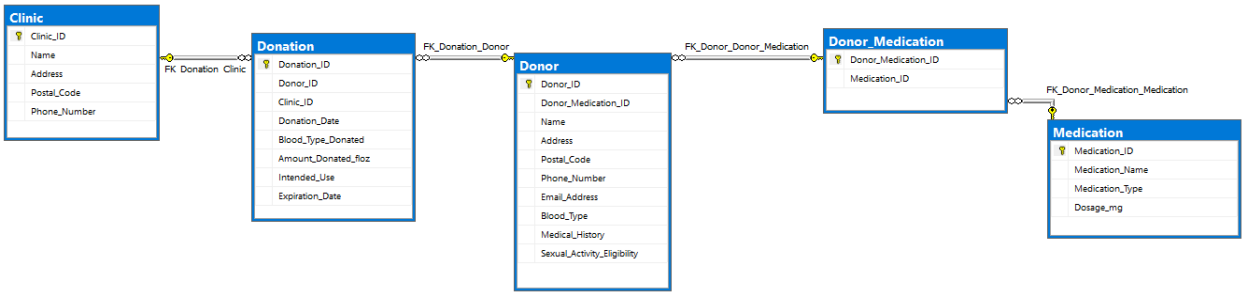

Figure 4 – The Schema Diagram for the database in MS SQL Server

SQL Code – Load Tables in SQL Server

- Randomized data was generated according to this database schema using Mockaroo.com.
- The randomized tables were saved as .csv files to a local machine.
- Data was added to the database using MS SQL Server's Import Data feature.
- 10 records were added to the Clinic table and 1000 records were added to the other tables.
- Since the scripts for the tables with 1000 records are large, only the script inserting data into the Clinic table is shown.
- The Clinic, Donation, Donor, Donor_Medication, and Medication tables are shown populated in figures 5, 6, 7, and 8, respectively.

Script Generated from Inserting 10 Records into the Clinic Table

```
USE [MDS523_Barstow_Final_Project]

GO

/****** Object:  Table [dbo].[Clinic]    Script Date: 12/6/2022 11:19:16 AM ******/

SET ANSI_NULLS ON

GO

SET QUOTED_IDENTIFIER ON

GO

CREATE TABLE [dbo].[Clinic](

        [Clinic_ID] [char](4) NOT NULL,

        [Name] [varchar](50) NOT NULL,

        [Address] [varchar](50) NOT NULL,

        [Postal_Code] [char](5) NOT NULL,

        [Phone_Number] [varchar](12) NOT NULL,
```

```
 CONSTRAINT [PK_Clinic] PRIMARY KEY CLUSTERED

(

	[Clinic_ID] ASC

)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]

) ON [PRIMARY]

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'1236',
N'Barton-Berge', N'73 Graceland Street', N'72523', N'799-648-8681')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'2714',
N'Anderson and Walsh', N'18069 Havey Court', N'46895', N'571-203-7827')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'3287',
N'Walter LLC', N'58127 Myrtle Terrace', N'10971', N'631-233-9276')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'3934',
N'Watsica Inc', N'82090 Ridgeview Drive', N'79076', N'190-610-9402')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'5201',
N'Veum Donations', N'7 Elka Crossing', N'58137', N'548-499-3831')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'7282',
N'Mercy Hospital', N'3655 Jay Place', N'32434', N'468-203-3008')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'7623',
N'Goldwater Clinic', N'34 Spaight Terrace', N'38352', N'129-155-5183')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'7680',
N'Better Health Clinic', N'42050 Kenwood Trail', N'11095', N'456-687-0507')

GO
```

```
INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'7885', N'Bechtelar-Goldner', N'620 Dixon Hill', N'90472', N'791-896-1470')

GO

INSERT [dbo].[Clinic] ([Clinic_ID], [Name], [Address], [Postal_Code], [Phone_Number]) VALUES (N'8257', N'Northwoods Health', N'22563 Florence Alley', N'28019', N'708-540-9734')

GO
```

Results Messages

	Clinic_ID	Name	Address	Postal_Code	Phone_Number
1	1236	Barton-Berge	73 Graceland Street	72523	799-648-8681
2	2714	Anderson and Walsh	18069 Havey Court	46895	571-203-7827
3	3287	Walter LLC	58127 Myrtle Terrace	10971	631-233-9276
4	3934	Watsica Inc	82090 Ridgeview Drive	79076	190-610-9402
5	5201	Veum Donations	7 Elka Crossing	58137	548-499-3831
6	7282	Mercy Hospital	3655 Jay Place	32434	468-203-3008
7	7623	Goldwater Clinic	34 Spaight Terrace	38352	129-155-5183
8	7680	Better Health Clinic	42050 Kenwood Trail	11095	456-687-0507
9	7885	Bechtelar-Goldner	620 Dixon Hill	90472	791-896-1470
10	8257	Northwoods Health	22563 Florence Alley	28019	708-540-9734

Figure 5 – The populated Clinic Table in the Database

	Donor_Medication_ID	Medication_ID
1	1057451953	2170906949
2	1063669113	2852615833
3	1082152726	6874343570
4	1083675803	4993576931
5	1085038110	9378423262
6	1085139354	5612947594
7	1086634701	7400058128
8	1086678563	3483151356
9	1098535268	8904340106
10	1099551708	2907072349

Figure 6 – The first 10 rows of the Donor_Medication Table in the Database

	Medication_ID	Medication_Name	Medication_Type	Dosage_mg
1	1002209730	Potassium Chloride in Dextrose	Analgesic	895
2	1002512973	Potassium Citrate	Anesthetic	665
3	1024124502	Cyclobenzaprine Hydrochloride	Antipsychotic	70
4	1024842608	PEMF Optimizer	Antidepressant	980
5	1027061495	Wal-Dryl	Tranquilizer	225
6	1048983202	Enoxaparin Sodium	Tranquilizer	290
7	1059680607	Quality Choice Vaginal Anti-Itch	Sedative	675
8	1078810510	ARRID EXTRA DRY	Analgesic	140
9	1079114511	Epivir	Mood Stabilizer	45
10	1084924971	CVS daily moisturizing	Tranquilizer	230

Figure 7 – The first 10 rows of the Medication Table in the Database

	Donor_ID	Donor_Medication_ID	Name	Address	Postal_Code	Phone_Number	Email_Address	Blood_Type	Medical_History	Sexual_Activity_Eligibility
1	1018806239	2589970727	Mackenzie Pol	2948 Delladonna Alley	59376	191-651-3090	mpolm2@theguardian.com	A-	Salmonellosis 2017	1
2	1020419285	3815155028	Meggy Folomkin	07137 Thierer Drive	88594	585-730-7517	mfolomkin12@163.com	O+		1
3	1042010648	6322038712	Melissa Vasile	81 Gateway Alley	74494	730-945-1452	mvasilebw@gravatar.com	O-		1
4	1042698713	8779848732	Waiter Darridon	0802 Troy Point	78835	627-639-7195	wdarridonfb@unblog.fr	O-		1
5	1049482632	7638924796	Phaidra Heddon	861 Namekagon Park	14801	760-375-7778	pheddon2z@hexun.com	A+		1
6	1057195024	9924190646	Thaddeus Pull	60 Eastwood Trail	81923	228-921-2300	tpull2n@usgs.gov	B-		0
7	1078952239	8225097245	Milly McGaugey	54 Annamark Avenue	23889	231-565-8861	mmcgaugey9t@mac.com	B+		0
8	1086663681	6386138217	Christiana Asson	00 Morningstar Crossing	58926	941-962-9786	cassono7@clickbank.net	B+		0
9	1090427030	9514830617	Ulrikaumeko Girke	3690 Corry Hill	99096	544-531-9228	ugirkemt@nytimes.com	O-		1
10	1102120827	4268859279	Shandra Carwithen	070 Surrey Junction	64311	374-436-0884	scarwithennk@nba.com	B+	Flu 2019	1

Figure 8 – The first 10 rows of the Donor Table in the Database

	Donation_ID	Donor_ID	Clinic_ID	Donation_Date	Blood_Type_Donated	Amount_Donated_floz	Intended_Use	Expiration_Date
1	1002997247	9572439136	8257	2022-04-17	B-	10.9	Platelets	2021-11-10
2	1006257909	7053119275	7885	2022-06-02	O-	5.1	RBC	2022-12-08
3	1010540374	2883853946	7680	2021-09-01	AB-	13.9	Platelets	2022-12-27
4	1011461915	1299213281	7680	2021-09-02	O+	7.5	Plasma	2023-09-04
5	1013410794	5801891573	7623	2021-12-28	AB-	12.8	Plasma	2022-09-03
6	1020579139	3808994655	2714	2021-07-14	O-	4.5	RBC	2022-11-17
7	1026308237	4933094415	7885	2022-04-01	AB-	14	Plasma	2022-07-20
8	1027431387	3574545218	7885	2021-05-19	B-	12.2	RBC	2022-10-06
9	1041754241	7557603285	1236	2021-04-18	AB-	14.4	Plasma	2021-07-14
10	1072502187	2155321712	7623	2022-04-19	O+	13.7	RBC	2021-08-05

Figure 9 – The first 10 rows of the Donation Table in the Database

Section 4 – Database Queries

Simple Queries

Query 1: How much of each blood type is there in inventory (in fluid ounces)?

```
SELECT Blood_Type_Donated, SUM(Amount_Donated_floz) AS Amount_Donated_floz
FROM Donation
GROUP BY Blood_Type_Donated
ORDER BY Blood_Type_Donated;
```

	Blood_Type_Donated	Amount_Donated_floz
1	A+	1145
2	A-	1113.2
3	AB+	992.7
4	AB-	1157
5	B+	1296.2
6	B-	1095.7
7	O+	859
8	O-	967.3

Figure 10 – Result of the first query. B+ blood is most common at 1296.2 fl. oz., O+ is the least common at 859 fl. oz.

Query 2: What is the breakdown of intended uses for the blood in inventory?

```
SELECT Intended_Use, COUNT(Intended_Use) AS Count_of_Intended_Use
FROM Donation
GROUP BY Intended_Use;
```

	Intended_Use	Count_of_Intended_Use
1	Platelets	349
2	RBC	314
3	Plasma	337

Figure 11 – Result of the second query. The most common intended use is platelets, the least is red blood cells.

Query 3: Which date had the largest number of donations?

```
SELECT Donation_Date, COUNT(Donation_Date) AS Count_of_Donation_Date
FROM Donation
GROUP BY Donation_Date
ORDER BY Count_of_Donation_Date DESC
OFFSET 0 ROWS FETCH NEXT 1 ROWS ONLY;
```

	Donation_Date	Count_of_Donation_Date
1	2022-03-20	8

Figure 12 – Result of the third query. The most was 8 donations on March, 20th 2022.

Query 4: Is there any blood past its expiry date in inventory (assuming today's date is 12/6/22)?

```
SELECT COUNT(Expiration_Date) AS Num_Expired
FROM Donation
WHERE Expiration_Date <= '2022-12-06';
```

	Num_Expired
1	649

Figure 13 – Result of the fourth query. More than half the inventory consists of expired blood.

Query 5: What are the phone numbers and email addresses of people with a certain blood type (say O-)?

```
SELECT Phone_Number, Email_Address, Blood_Type
FROM Donor
WHERE Blood_Type = 'O-';
```

	Phone_Number	Email_Address	Blood_Type
1	730-945-1452	mvasilebw@gravatar.com	O-
2	627-639-7195	wdarridonfb@unblog.fr	O-
3	544-531-9228	ugirkemt@nytimes.com	O-
4	937-734-7654	cmeffinw@weebly.com	O-
5	701-823-1619	bhearneis@over-blog.com	O-
6	623-137-0482	cloundp6@youtube.com	O-
7	536-921-2477	nbalmann2p@behance.net	O-
8	987-577-5671	sskamallu@ebay.co.uk	O-
9	568-404-4008	brojel2@squidoo.com	O-
10	695-670-9465	astride6p@marketwatch.com	O-
11	197-516-6227	chubboldln@list-manage.com	O-
12	653-557-4158	jleving9o@mozilla.com	O-
13	904-228-2249	dingliskd@newsvine.com	O-
14	968-570-6790	aclouttrr@sogou.com	O-
15	368-592-4380	jrait9k@sun.com	O-
16	905-753-6139	ycrowncm@wordpress.com	O-
17	296-847-6902	pchrippesap@hostgator.com	O-
18	347-703-8283	pdochartycp@wikispaces.com	O-
19	616-436-9892	psayerpb@deviantart.com	O-

Figure 14 – Result of the fifth query. There were 127 people with O- blood.

Complex Queries

Query 6: What is the name of the clinic that has the most of a certain blood type on-hand (say AB+)?

```
SELECT  d.Blood_Type_Donated, d.Clinic_ID, c.Name, SUM(d.Amount_Donated_floz)
AS Amount_Donated_floz
FROM Donation d
INNER JOIN Clinic c
ON d.Clinic_ID = c.Clinic_ID
WHERE d.Blood_Type_Donated = 'AB+'
GROUP BY c.Name, d.Clinic_ID, d.Blood_Type_Donated
ORDER BY Amount_Donated_floz DESC;
```

	Blood_Type_Donated	Clinic_ID	Name	Amount_Donated_floz
1	AB+	8257	Northwoods Health	164.9
2	AB+	5201	Veum Donations	126.6
3	AB+	3287	Walter LLC	121.1
4	AB+	3934	Watsica Inc	121.1
5	AB+	7623	Goldwater Clinic	111.3
6	AB+	1236	Barton-Berge	102
7	AB+	7282	Mercy Hospital	96.8
8	AB+	7885	Bechtelar-Goldner	80.1
9	AB+	7680	Better Health Clinic	39.9
10	AB+	2714	Anderson and Walsh	28.9

Figure 15 – The result of the sixth query. Northwoods health collected 164.9 fl. oz. of AB+ blood.

Query 7: What are the names of medications a particular donor takes (Dina Canon)?

```
SELECT d.Donor_Medication_ID AS d_ID, dm.Donor_Medication_ID AS dm_ID, d.Name,
m.Medication_Name, m.Medication_Type, m.Dosage_mg
FROM Donor d
INNER JOIN Donor_Medication dm ON d.Donor_Medication_ID = dm.Donor_Medication_ID
INNER JOIN Medication m ON dm.Medication_ID = m.Medication_ID
WHERE d.name = 'Dina Canon';
```

	d_ID	dm_ID	Name	Medication_Name	Medication_Type	Dosage_mg
1	2394049636	2394049636	Dina Canon	Oxycodone and Acetaminophen	Antipsychotic	475

Figure 16 – The result of the seventh query. Dina Canon takes a 475mg combination of Oxycodone and Acetaminophen.

Query 8: What are the phone numbers of people who take a certain medication (say Epivir, Trexall, and Oxacillin)?

```
SELECT d.Phone_Number, d.Name, m.Medication_Name, m.Dosage_mg, d.Donor_Medication_ID AS
d_ID, dm.Donor_Medication_ID as dm_ID
FROM Donor d
INNER JOIN Donor_Medication dm ON d.Donor_Medication_ID = dm.Donor_Medication_ID
INNER JOIN Medication m ON dm.Medication_ID = m.Medication_ID
WHERE m.Medication_Name = 'Epivir'
OR m.Medication_Name = 'Trexall'
OR m.Medication_Name = 'Oxacillin';
```

	Phone_Number	Name	Medication_Name	Dosage_mg	d_ID	dm_ID
1	680-376-6841	Joete Gittus	Trexall	650	9391738676	9391738676
2	761-221-2654	Myrwyn Trammel	Oxacillin	260	7138956133	7138956133
3	838-186-6059	Oriana Rosthorn	Trexall	650	5581338624	5581338624
4	990-490-1233	Randolph Hendricks	Oxacillin	260	3138849391	3138849391
5	424-222-3835	Claudine Bramich	Oxacillin	260	8413537021	8413537021

Figure 17 – The result of the eighth query. Evidently, no takes Epivir.

Query 9: What is/are the name(s) of the clinic(s) that took donations on the most recent date in the database (7/31/22)?

```
SELECT c.Clinic_ID AS c_ID, d.Clinic_ID AS d_ID, d.Donation_Date, c.Name, c.Address,
c.Postal_Code, c.Phone_Number
FROM Clinic c
INNER JOIN Donation d ON c.Clinic_ID = d.Clinic_ID
WHERE d.Donation_Date = '2022-07-31';
```

	c_ID	d_ID	Donation_Date	Name	Address	Postal_Code	Phone_Number
1	7885	7885	2022-07-31	Bechtelar-Goldner	620 Dixon Hill	90472	791-896-1470
2	8257	8257	2022-07-31	Northwoods Health	22563 Florence Alley	28019	708-540-9734

Figure 18 – The result of the ninth query. There were two clinics that took donations on the most recent date in the database: Bechtelar-Goldner and Northwoods Health.

Query 10: What are all the names of clinics that a particular donor has donated at before (Aeriel Durie)?

```
SELECT c.Name, dr.Name, dn.Donation_Date
FROM Clinic c
INNER JOIN Donation dn ON c.Clinic_ID = dn.Clinic_ID
INNER JOIN Donor dr ON dn.Donor_ID = dr.Donor_ID
WHERE dr.Name = 'Aeriel Durie';
```

	Name	Name	Donation_Date
1	Northwoods Health	Aeriel Durie	2021-06-26
2	Goldwater Clinic	Aeriel Durie	2021-10-09

Figure 19 – The result of the tenth query. Aeriel has donated at Northwoods Health Goldwater Clinic.

Section 5 – Data Warehouse Design and Coding

Data Warehouse Design

- This small data warehouse has 1 fact table, Clinic Fact, with 2 dimension tables, Donation Dimension and Clinic Dimension.
- The dimensional model for the data warehouse is shown in fig. 20.
- Although both a Donation Fact table and a Date Dimension were proposed, they were not included. These could easily be added at a later time.

There were 3 questions this data warehouse sought to answer:

1) Which clinics had the most/least donations overall?
2) Which blood types does each clinic need the most?
3) What clinic most needs a particular blood type?

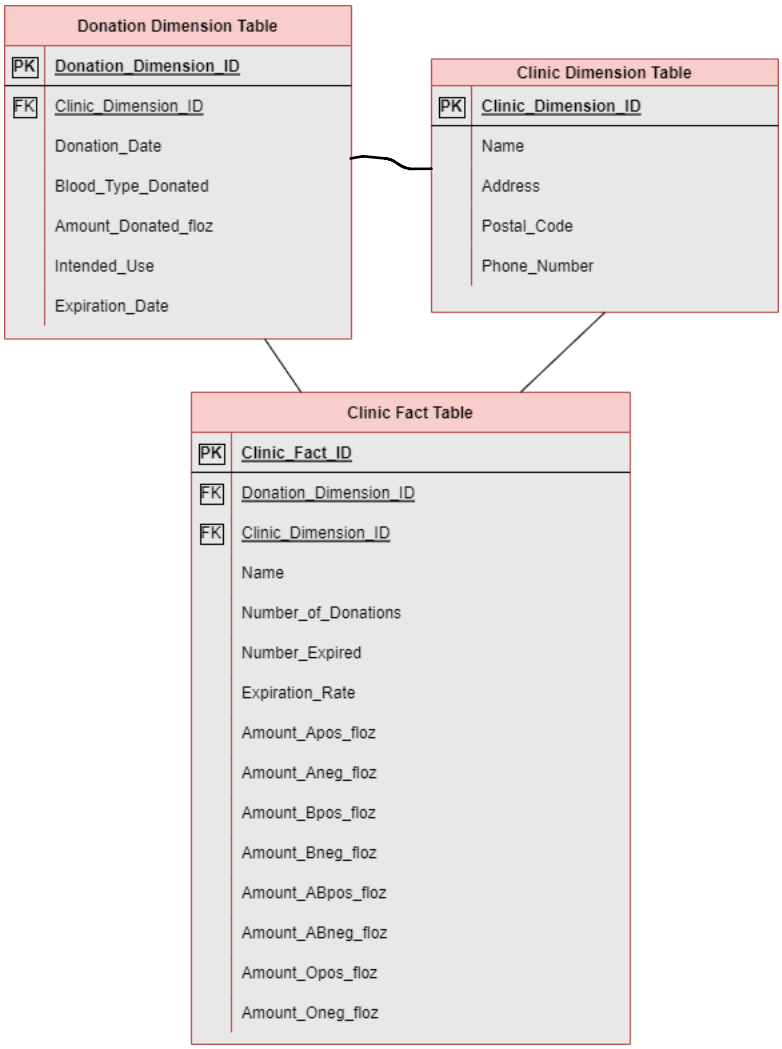

Figure 20 – The Dimensional Model for the Data Warehouse

Data Marts

- If the data warehouse was larger like in fig. 21, then an appropriate 3-table data mart would be the subset: Donation Fact Table, Date Dimension Fact Table, and Donation Dimension Fact Table.
- The data mart would allow for week-to-week or month-to-month trends.

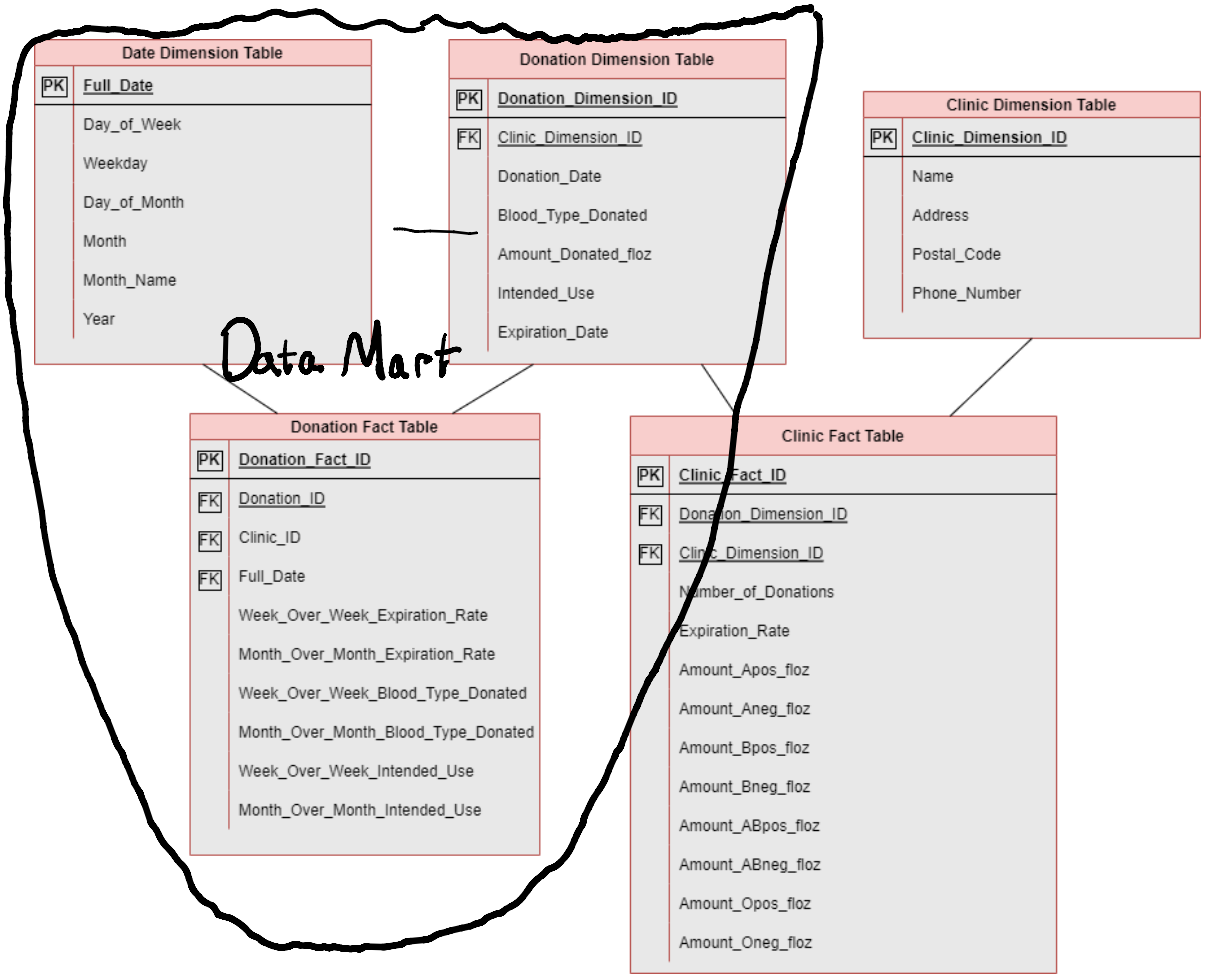

Figure 21 – The possible 3-table data mart for a data warehouse larger than the one created in this project.

Data Warehouse Creation – Code

- The Clinic Dimension table in the data warehouse is essentially the same as the Clinic table in the database. The primary key is the same as the Clinic_ID in the database.
- The Donation Dimension table in the data warehouse is basically a select few columns from the Donation table in the database. The primary key is the same as the Donation_ID in the database.
- The Clinic Fact table consists of summary information and the rolled-up totals for each blood type for each clinic. The primary key is an auto-incrementing integer.

SQL Code - Scripts to Create Data Warehouse Fact and Dimension Tables in SQL Server

```
USE [MDS523_Barstow_Final_Project_DW]
GO
/****** Object:  Table [dbo].[Clinic_Dimension]    Script Date: 12/8/2022 1:50:32 AM ******/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[Clinic_Dimension](
	[Clinic_Dimension_ID] [char](4) NOT NULL,
	[Name] [varchar](50) NOT NULL,
	[Address] [varchar](50) NOT NULL,
	[Postal_Code] [char](5) NOT NULL,
	[Phone_Number] [varchar](12) NOT NULL,
 CONSTRAINT [PK_Clinic_Dimension] PRIMARY KEY CLUSTERED
(
	[Clinic_Dimension_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
USE [MDS523_Barstow_Final_Project_DW]
GO
/****** Object:  Table [dbo].[Clinic_Fact]    Script Date: 12/8/2022 1:51:05 AM ******/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
```

```
CREATE TABLE [dbo].[Clinic_Fact](
	[Clinic_Fact_ID] [int] IDENTITY(1,1) NOT NULL,
	[Donation_Dimension_ID] [char](10) NOT NULL,
	[Clinic_Dimension_ID] [char](4) NOT NULL,
	[Name] [varchar](50) NOT NULL,
	[Number_Of_Donations] [int] NOT NULL,
	[Expiration_Rate] [float] NOT NULL,
	[Amount_Apos_floz] [float] NOT NULL,
	[Amount_Aneg_floz] [float] NOT NULL,
	[Amount_Bpos_floz] [float] NOT NULL,
	[Amount_Bneg_floz] [float] NOT NULL,
	[Amount_ABpos_floz] [float] NOT NULL,
	[Amount_ABneg_floz] [float] NOT NULL,
	[Amount_Opos_floz] [float] NOT NULL,
	[Amount_Oneg_floz] [float] NOT NULL,
 CONSTRAINT [PK_Clinic_Fact] PRIMARY KEY CLUSTERED
(
	[Clinic_Fact_ID] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF,
ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
ALTER TABLE [dbo].[Clinic_Fact]  WITH CHECK ADD  CONSTRAINT [FK_Clinic_Fact_Clinic_Dimension]
FOREIGN KEY([Clinic_Dimension_ID])
REFERENCES [dbo].[Clinic_Dimension] ([Clinic_Dimension_ID])
GO
ALTER TABLE [dbo].[Clinic_Fact] CHECK CONSTRAINT [FK_Clinic_Fact_Clinic_Dimension]
GO
```

```
ALTER TABLE [dbo].[Clinic_Fact]  WITH CHECK ADD  CONSTRAINT [FK_Clinic_Fact_Donation_Dimension] FOREIGN KEY([Donation_Dimension_ID])

REFERENCES [dbo].[Donation_Dimension] ([Donation_Dimension_ID])

GO

ALTER TABLE [dbo].[Clinic_Fact] CHECK CONSTRAINT [FK_Clinic_Fact_Donation_Dimension]

GO

USE [MDS523_Barstow_Final_Project_DW]

GO

/****** Object:  Table [dbo].[Donation_Dimension]    Script Date: 12/8/2022 1:51:43 AM ******/

SET ANSI_NULLS ON

GO

SET QUOTED_IDENTIFIER ON

GO

CREATE TABLE [dbo].[Donation_Dimension](

        [Donation_Dimension_ID] [char](10) NOT NULL,

        [Clinic_Dimension_ID] [char](4) NOT NULL,

        [Donation_Date] [date] NOT NULL,

        [Blood_Type_Donated] [varchar](3) NOT NULL,

        [Amount_Donated_floz] [float] NOT NULL,

        [Intended_Use] [varchar](9) NOT NULL,

        [Expiration_Date] [date] NOT NULL,

 CONSTRAINT [PK_Donation_Dimension] PRIMARY KEY CLUSTERED

(

        [Donation_Dimension_ID] ASC

)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]

) ON [PRIMARY]

GO
```

ALTER TABLE [dbo].[Donation_Dimension] WITH CHECK ADD CONSTRAINT [FK_Donation_Dimension_Clinic_Dimension] FOREIGN KEY([Clinic_Dimension_ID])

REFERENCES [dbo].[Clinic_Dimension] ([Clinic_Dimension_ID])

GO

ALTER TABLE [dbo].[Donation_Dimension] CHECK CONSTRAINT [FK_Donation_Dimension_Clinic_Dimension]

GO

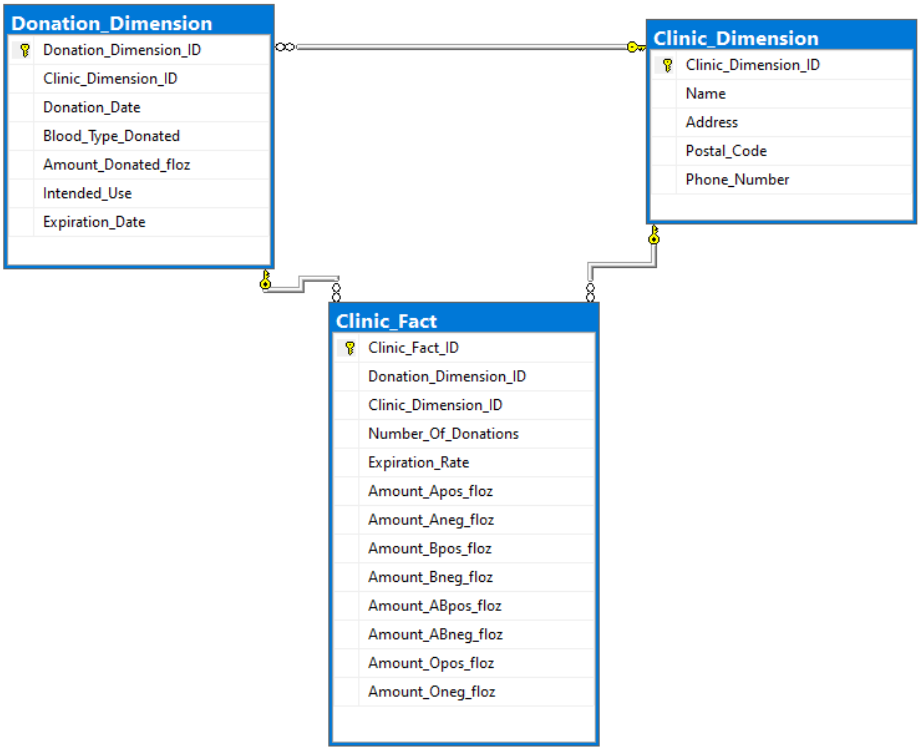

Figure 22 – The Schema Diagram for the data warehouse in MS SQL Server

SQL Code – Script to Load Data into the 3 Tables in the Data Warehouse in SQL Server

```
INSERT INTO Clinic_Dimension (Clinic_Dimension_ID, Name, Address, Postal_Code,
Phone_Number)
	SELECT c.Clinic_ID AS Clinic_Dimension_ID,
	c.Name AS Name,
	c.Address AS Address,
	c.Postal_Code AS Postal_Code,
	c.Phone_Number AS Phone_Number
	FROM MDS523_Barstow_Final_Project_DB.dbo.clinic AS c
```

```
INSERT INTO Donation_Dimension (Donation_Dimension_ID, Clinic_Dimension_ID,
Donation_Date, Blood_Type_Donated, Amount_Donated_floz, Intended_Use, Expiration_Date)
       SELECT d.Donation_ID AS Donation_Dimension_ID,
       d.Clinic_ID AS Clinic_Dimension_ID,
       d.Donation_Date AS Donation_Date,
       d.Blood_Type_Donated AS Blood_Type_Donated,
       d.Amount_Donated_floz AS Amount_Donated_floz,
       d.Intended_Use AS Intended_Use,
       d.Expiration_Date AS Expiration_Date
       FROM MDS523_Barstow_Final_Project_DB.dbo.donation AS d


/*Inserting clinic name and ID*/
INSERT INTO Clinic_Fact (Clinic_Dimension_ID, Name)
       SELECT c.Clinic_ID as Clinic_Dimension_ID,
       c.Name as Name
       FROM MDS523_Barstow_Final_Project_DB.dbo.Clinic AS c

/*Updating number of donations*/
UPDATE Clinic_Fact
       SET Number_Of_Donations =
       (SELECT COUNT(d.Clinic_ID)
       FROM MDS523_Barstow_Final_Project_DB.dbo.donation AS d
       WHERE d.Clinic_ID = Clinic_Dimension_ID)

/*Updating number expired*/
UPDATE Clinic_Fact
       SET Number_Expired =
       (SELECT COUNT(d.Expiration_Date)
       FROM MDS523_Barstow_Final_Project_DB.dbo.donation AS d
       WHERE d.Clinic_ID = Clinic_Dimension_ID
       AND
       d.Expiration_Date <= '2022-12-06')

/*Updating expiration rate*/
UPDATE Clinic_Fact
       SET Expiration_Rate =
       (SELECT Number_Expired/Number_Of_Donations
       FROM Clinic_Fact);

/*Updating each blood type*/
UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'A+')

UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'A-')
```

```
UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'B+')

UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'B-')

UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'AB+')

UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'AB-')

UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'O+')

UPDATE Clinic_Fact
       SET Amount_Apos_floz =
       (SELECT SUM(d.Amount_Donated_floz)
       FROM Donation_Dimension d
       INNER JOIN MDS523_Barstow_Final_Project_DB.dbo.clinic c
       ON d.Clinic_Dimension_ID = Clinic_Dimension_ID
       WHERE d.Blood_Type_Donated = 'O-')
```

	Donation_Dimension_ID	Clinic_Dimension_ID	Donation_Date	Blood_Type_Donated	Amount_Donated_floz	Intended_Use	Expiration_Date
1	1002997247	8257	2022-04-17	B-	10.9	Platelets	2021-11-10
2	1006257909	7885	2022-06-02	O-	5.1	RBC	2022-12-08
3	1010540374	7680	2021-09-01	AB-	13.9	Platelets	2022-12-27
4	1011461915	7680	2021-09-02	O+	7.5	Plasma	2023-09-04
5	1013410794	7623	2021-12-28	AB-	12.8	Plasma	2022-09-03
6	1020579139	2714	2021-07-14	O-	4.5	RBC	2022-11-17
7	1026308237	7885	2022-04-01	AB-	14	Plasma	2022-07-20
8	1027431387	7885	2021-05-19	B-	12.2	RBC	2022-10-06
9	1041754241	1236	2021-04-18	AB-	14.4	Plasma	2021-07-14
10	1072502187	7623	2022-04-19	O+	13.7	RBC	2021-08-05

Figure 23 – The first 10 rows of the Donation Dimension table in the data warehouse

	Clinic_Dimension_ID	Name	Address	Postal_Code	Phone_Number
1	1236	Barton-Berge	73 Graceland Street	72523	799-648-8681
2	2714	Anderson and Walsh	18069 Havey Court	46895	571-203-7827
3	3287	Walter LLC	58127 Myrtle Terrace	10971	631-233-9276
4	3934	Watsica Inc	82090 Ridgeview Drive	79076	190-610-9402
5	5201	Veum Donations	7 Elka Crossing	58137	548-499-3831
6	7282	Mercy Hospital	3655 Jay Place	32434	468-203-3008
7	7623	Goldwater Clinic	34 Spaight Terrace	38352	129-155-5183
8	7680	Better Health Clinic	42050 Kenwood Trail	11095	456-687-0507
9	7885	Bechtelar-Goldner	620 Dixon Hill	90472	791-896-1470
10	8257	Northwoods Health	22563 Florence Alley	28019	708-540-9734

Figure 24 – The Clinic Dimension table in the data warehouse

	Clinic_Fact_ID	Donation_Dimension_ID	Clinic_Dimension_ID	Name	Number_Of_Donations	Number_Expired	Expiration_Rate	Amount_Apos_floz	Amount_Aneg_floz	Amount_Bpos_floz	Amount_Bneg_floz	Amount_ABpos_floz	Amount_ABneg_floz	Amount_Opos_floz	Amount_Oneg_floz
1	1	NULL	1236	Barton-Berge	100	66	0.66	145.4	134.2	137.6	127.8	28.9	111.3	114.3	131
2	2	NULL	2714	Anderson and Walsh	100	64	0.64	78.8	108.9	143.9	133.5	102	142.4	57	62.5
3	3	NULL	3287	Walter LLC	100	56	0.56	151	128.5	112.4	68.7	80.1	90.5	53.9	60.3
4	4	NULL	3934	Watsica Inc	100	68	0.68	102.3	198.3	109.9	86.3	39.9	131.2	82.9	162
5	5	NULL	5201	Veum Donations	100	55	0.55	156.3	101.8	100.4	57.4	111.3	135.2	135.8	46.8
6	6	NULL	7282	Mercy Hospital	100	65	0.65	123.7	109	97.4	66.9	96.8	151.9	92.9	98.3
7	7	NULL	7623	Goldwater Clinic	100	67	0.67	81.8	55.3	163.3	155.2	164.9	93.8	105.9	91.6
8	8	NULL	7680	Better Health Clinic	100	67	0.67	95.4	68.1	228.5	126.6	126.6	53.5	93.3	70.8
9	9	NULL	7885	Bechtelar-Goldner	100	70	0.7	97.5	110	73.8	158.5	121.1	102.6	76.7	147.4
10	10	NULL	8257	Northwoods Health	100	71	0.71	112.8	99.1	129	114.8	121.1	144.6	46.3	96.6

Figure 25 – The Clinic Fact table in the data warehouse

Data Warehouse Queries – Code

Query 1: Which clinics had the most/least donations overall?

```
SELECT Name, Number_Of_Donations
FROM Clinic_Fact;
```

	Name	Number_Of_Donations
1	Barton-Berge	100
2	Anderson and Walsh	100
3	Walter LLC	100
4	Watsica Inc	100
5	Veum Donations	100
6	Mercy Hospital	100
7	Goldwater Clinic	100
8	Better Health Clinic	100
9	Bechtelar-Goldner	100
10	Northwoods Health	100

Figure 26 – The result of the first query. As a result of distributing the random data uniformly, each clinic has collected 100 donations.

Query 2: What is the expiration rate for each clinic?

```
SELECT Name, Expiration_Rate
FROM Clinic_Fact
ORDER BY Expiration_Rate;
```

	Name	Expiration_Rate
1	Veum Donations	0.55
2	Walter LLC	0.56
3	Anderson and Walsh	0.64
4	Mercy Hospital	0.65
5	Barton-Berge	0.66
6	Goldwater Clinic	0.67
7	Better Health Clinic	0.67
8	Watsica Inc	0.68
9	Bechtelar-Goldner	0.7
10	Northwoods Health	0.71

Figure 27 – The result of the second query. Veum Donations has the lowest expiration rate at 55%, while Northwoods Health has the highest at 71%

Query 3: Which clinic most needs a particular blood type (say B-)?

	Name	Amount_Bneg_floz
1	Veum Donations	57.4
2	Mercy Hospital	66.9
3	Walter LLC	68.7
4	Watsica Inc	86.3
5	Northwoods Health	114.8
6	Better Health Clinic	126.6
7	Barton-Berge	127.8
8	Anderson and Walsh	133.5
9	Goldwater Clinic	155.2
10	Bechtelar-Goldner	158.5

Figure 28 – The result of the third query. Veum Donations has the fewest fl. oz. of B- blood in their inventory, so they need B- blood the most.

Conclusion

This project looked at creating a data warehouse for a small blood donation business. The data warehouse allows the business to analyze the breakdown of blood types at each clinic and to look at the expiration rates of each clinic. This information will help the business augment their operations, reducing their expiration rates by reallocating the blood types available in each clinic's inventory. Further suggestions to improve the data warehouse would be to include Donor and Date Dimension tables and Donor and Donation fact tables. Doing this could allow for time series analysis, looking at week-to-week breakdowns of each blood type collected. It could also allow for more robust analytics about the Donors themselves, such as whether someone who has a needed blood type has been denied donating because of sexual activity. Ultimately, the database will serve the business's operational needs, and the data warehouse will serve its analytical needs to drive improvement.

Acknowledgements

I would like to thank Dr. Jensen for his feedback on the designs of the database and data warehouse and for his help during the diagramming phases. I also want to thank Jim Wheeler for his input on database design in general.

References

[1] https://donatebloodnow.org/what-it-takes-to-start-a-blood-donation-center/

[2] https://www.sharecare.com/health/blood-basics/how-does-the-blood-donation-process-work#:~:text=A%20health%20history%20and%20mini,blood%20cells%2C%20plasma%20and%20platelets.

[3] https://www.redcrossblood.org/donate-blood/blood-donation-process/donation-process-overview.html

[4] https://www.redcrossblood.org/donate-blood/blood-donation-process/what-happens-to-donated-blood.html

[5] https://www.webmd.com/a-to-z-guides/what-to-expect-when-you-give-blood

[6] https://www.redcrossblood.org/donate-blood/how-to-donate/eligibility-requirements/eligibility-criteria-alphabetical.html