

Telecommunications Churn Analysis

Author: Beth Fuller

Date: January 2024

A: Question

What type of internet is being provided to customers that both did and did not churn?

A1: Identifying Data

The data that will be used for this analysis will first come from the churn database. This database contains information from 10,000 customers. Within the database, there are several tables. This analysis will use the customer's table. Within the customer table, we will use the following columns:

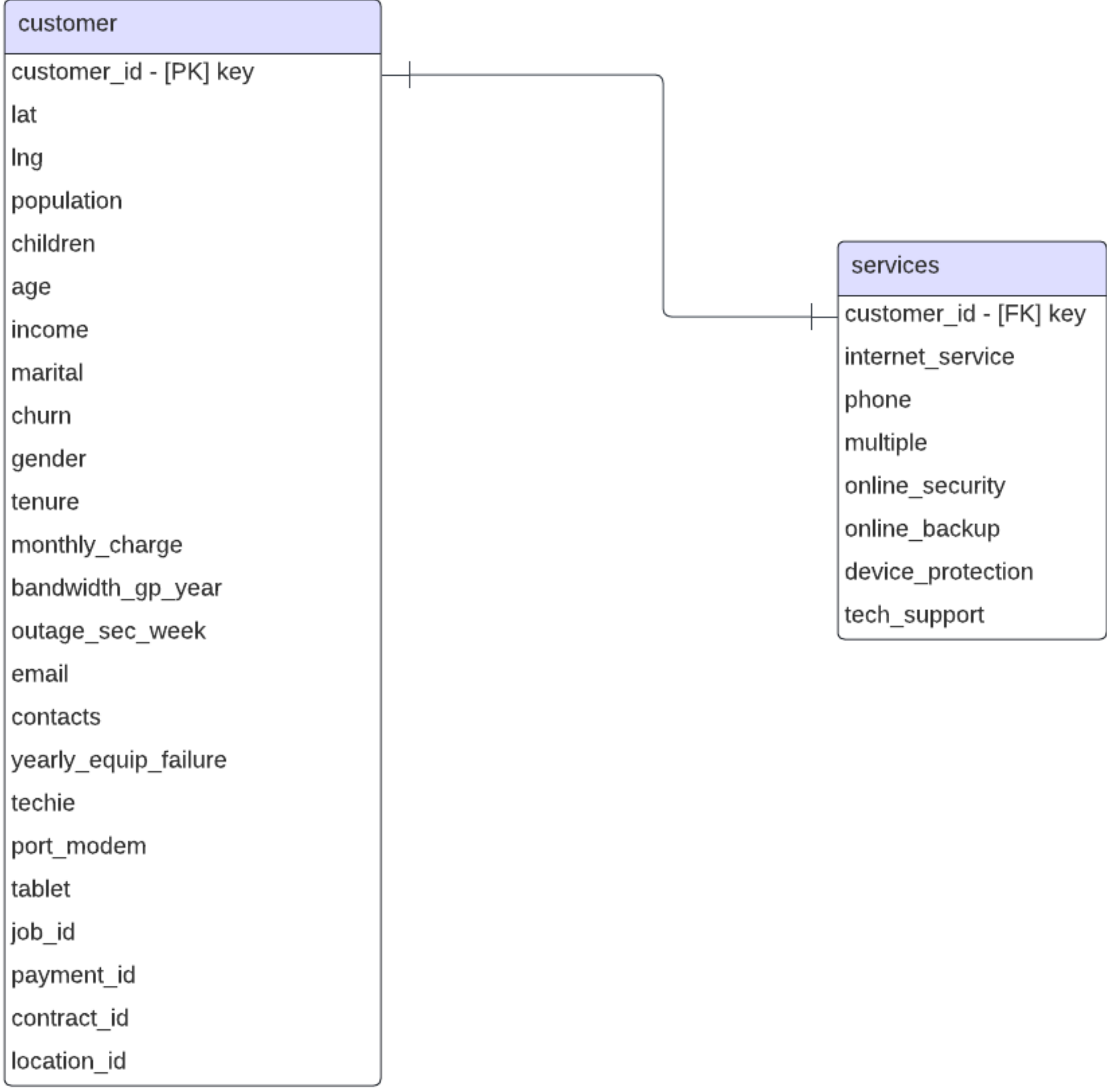
- customer_id - [PK] text
- churn - text
- monthly_charge - numeric

The second source of data will be the services.csv file that contains customer IDs and 7 questions about services used by each customer. Once the table is created from this csv file the following columns will be used:

- customer_id - [PK] text
- internet_service - text
- phone - boolean

B: Entity Relationship Diagram

The entity relationship diagram is shown below. The customer and services table are connected via the primary key customer_id. There is a one-to-one relationship ("What is an Entity Relationship Diagram (ERD)?", 2024).



B1: ERD Code

```
In [ ]: BEGIN;

CREATE TABLE IF NOT EXISTS public.customer
(
    customer_id text COLLATE pg_catalog."default" NOT NULL,
    lat text COLLATE pg_catalog."default",
    lng text COLLATE pg_catalog."default",
    population text COLLATE pg_catalog."default",
    children text COLLATE pg_catalog."default",
    age text COLLATE pg_catalog."default",
    income text COLLATE pg_catalog."default",
    marital text COLLATE pg_catalog."default",
    churn text COLLATE pg_catalog."default",
    gender text COLLATE pg_catalog."default",
    tenure text COLLATE pg_catalog."default",
    monthly_charge text COLLATE pg_catalog."default",
    bandwidth_gp_year text COLLATE pg_catalog."default",
    outage_sec_week text COLLATE pg_catalog."default",
    email text COLLATE pg_catalog."default",
    contacts text COLLATE pg_catalog."default",
    yearly_equip_failure text COLLATE pg_catalog."default",
    techie text COLLATE pg_catalog."default",
    port_modem text COLLATE pg_catalog."default",
    tablet text COLLATE pg_catalog."default",
    job_id text COLLATE pg_catalog."default",
    payment_id text COLLATE pg_catalog."default",
    contract_id text COLLATE pg_catalog."default",
    location_id text COLLATE pg_catalog."default",
    CONSTRAINT customer_pkey PRIMARY KEY (customer_id)
);

CREATE TABLE IF NOT EXISTS public.services
(
    customer_id text COLLATE pg_catalog."default" NOT NULL,
    internet_service text COLLATE pg_catalog."default" NOT NULL,
    phone boolean NOT NULL,
    multiple boolean NOT NULL,
    online_security boolean NOT NULL,
    online_backup boolean NOT NULL,
    device_protection boolean NOT NULL,
    tech_support boolean NOT NULL,
    CONSTRAINT "services_pkey" PRIMARY KEY (customer_id)
);

ALTER TABLE IF EXISTS public.services
    ADD FOREIGN KEY (customer_id)
    REFERENCES public.customer (customer_id) MATCH SIMPLE
    ON UPDATE NO ACTION
    ON DELETE NO ACTION
    NOT VALID;

END;
```

B1: Create Service Table

```
In [ ]: CREATE TABLE public.services
(
    customer_id text COLLATE pg_catalog."default" NOT NULL,
    internet_service text COLLATE pg_catalog."default" NOT NULL,
    phone boolean NOT NULL,
    multiple boolean NOT NULL,
    online_security boolean NOT NULL,
    online_backup boolean NOT NULL,
    device_protection boolean NOT NULL,
    tech_support boolean NOT NULL,
    CONSTRAINT "services_pkey" PRIMARY KEY (customer_id)
)

TABLESPACE pg_default;

ALTER TABLE public.services
    OWNER to postgres;
```

B2: Load Service Data

```
In [ ]: COPY public.services (customer_id, internet_service, phone, multiple, online_security, online_backup, device_protection, tech_support)
FROM 'C:/LabFiles/Services.csv'
DELIMITER ','
CSV HEADER QUOTE '"'
ESCAPE '\\';
```

C: SQL Query

What type of internet is being provided to customers that both did and did not churn?

This query shows that customers who are being lost have both DSL and Fiber Optic internet service options. It also shows customers that are continuing with service tend to utilize Fiber Optic internet service more than using DSL.

```
In [ ]: SELECT customer.churn, services.internet_service, COUNT(*)
FROM customer
INNER JOIN services ON customer.customer_id=services.customer_id
GROUP BY services.internet_service, customer.churn
ORDER BY churn DESC, services.internet_service DESC;
```

	churn text	internet_service text	count bigint
1	Yes	None	496
2	Yes	Fiber Optic	1040
3	Yes	DSL	1114
4	No	None	1633
5	No	Fiber Optic	3368
6	No	DSL	2349

D: Time / Database Refresh Guidelines

The company offers plans that are either month-to-month, one-year or two-year. The data should be updated on the first of each month to give insight in the customers being lost or gained each month.

E: Panopto Video of Code

<https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=14f676bb-ee17-4767-89a3-b0fc013d650d>

F: Web Sources

What is an Entity Relationship Diagram (ERD)?. (2024, January 10). Lucidchart. <https://www.lucidchart.com/pages/er-diagrams>

G: Sources

Western Governors University - D205. (2024). Scenario: Telecommunications Churn. <https://access.wgu.edu/ASP3/aap/content/g9rke9s0rlc9ejd92md0.html>

```
CREATE TABLE public.services
(
    customer_id text COLLATE pg_catalog."default" NOT NULL,
    internet_service text COLLATE pg_catalog."default" NOT NULL,
    phone boolean NOT NULL,
    multiple boolean NOT NULL,
    online_security boolean NOT NULL,
    online_backup boolean NOT NULL,
    device_protection boolean NOT NULL,
    tech_support boolean NOT NULL,
    CONSTRAINT "services _pkey" PRIMARY KEY (customer_id)
)

TABLESPACE pg_default;

ALTER TABLE public.services
    OWNER to postgres;
```

```
COPY public.services (customer_id, internet_service, phone, multiple,  
online_security, online_backup, device_protection, tech_support)  
FROM 'C:/LabFiles/Services.csv'  
DELIMITER ','  
CSV HEADER QUOTE ''''  
ESCAPE '''';
```

```
SELECT customer.churn, services.internet_service, COUNT(*)  
FROM customer  
INNER JOIN services ON customer.customer_id=services.customer_id  
GROUP BY services.internet_service, customer.churn  
ORDER BY churn DESC, services.internet_service DESC;
```