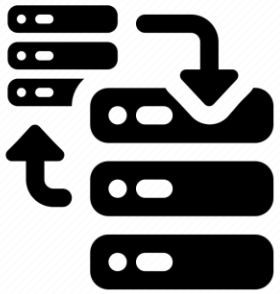


Data transformation & migration modeling for Practice Management

Electronic health record (EHR) & practice management software provider

- Partnered with the client to overcome challenges in migrating existing Practice Management data for Med Spas into their proprietary system via third-party partners
- Developed an efficient and economical automated solution to alleviate the time-consuming and costly nature of the existing migration process

Healthcare platform seeks a “seamless" cure for costly data transformation and migration

Picture this…

You’re looking to simplify the migration of Practice Management (PM) data, as current migration solutions are neither cost-effective nor streamlined for your customers’ needs

You turn to Accordion.

We partner with your team to develop an automated data transformation and migration solution to alleviate the time-consuming and costly nature of the existing migration process, including:

1) Building a single-click model using Python and Excel that simplifies the data transformation and ingestion process, saving time and resources
2) Providing Practice Management data outputs with >99% mapping between patients and appointments, enabling effective communication and patient reporting
3) Generating automated detailed exceptions report to identify and address data quality issues, ensuring only high-quality data is imported

Your value is enhanced.

- You have streamlined the Practice Management data migration process, reducing costs for your customers by 90% and decreasing data transformation and ingestion time by 95%, enabling faster onboarding
- You have also enhanced data accuracy by 15% to ensure high-quality PM data migration
- Additionally, you have eliminated the need for the technical team involvement in the end-to-end migration, freeing up resources and saving approximately 120 FTE-hours per month
- You have enabled the sales team to add $2.5M in annual revenue by enhancing their onboarding capacity

DATA TRANSFORMATION & MIGRATION MODELING

KEY RESULT

- 90% reduction in migration cost
- Potential annual revenue addition of $2.5M
- 95% faster ingestion time
- 15% improvement in data accuracy

VALUE LEVERS PULLED

- Migration modeling
- Exceptions reporting

Methodology/Approach

The approach for any EHR data transformation and migration model would consist of the five steps below:

01	02	03	04	05
Uploading input datasets	**Data cleaning & regex validations**	**User-defined inputs**	**Patient ID generation & mapping**	**Final output**
• Import the **patients** and **appointments** datasets in a predefined input format to maintain input consistency	• Clean the input data to **avoid invalid/inaccurate values** and maintain data sanctity • Carry out **regular expression validations** for each field value to comply with the output field formats • Provide visibility into the execution progress through **progress bars** to **improve user experience**	• Take inputs such as starting Patient ID, Duration, etc. from the user to create an output file that is **readily ingestible** into target system • Allow flexibility for users to export only **relevant data** in the output. For example, option to export the cancelled appointments or not.	• Use the **generated patient IDs** to **map** the input appointments with the input patients • Account for the **non-patient appointments** separately based on user inputs	• **Patients.csv:** Contains the patients' data as per output specifications • **Appointments.csv**: Contains the appointments data as per output specifications • **Exceptions & Summary.xlsx:** Contains all data transformations details and the reasons for purging specific rows/records during migration

Note : *All the models are designed to be case insensitive in terms of input column headers.*

Input patients and appointments data

CID	ClientName	Address	City	State	Postal	Email	CellPhone	Balance	Gender	Birthday	DateCreated
2011	ABC XYZ	Xytz Drive	Springdale	TX	33206	abc@mail.com	234-734-2321	$4960	M	5/5/1988	11/15/2016
2012	BCD WXY	Qwib Lane	Pinecrest	TY	86651	ag2@mail.com	917-232-2394	$55	F	6/30/1997	1/17/2014
2013	ABH BHJ	Zorn Avenue	Willow	CA	45314	ge2@mail.com	713-238-3231	$100	M	6/27/1989	8/28/2015
2014	ABH NJA	Fyx Street	Oakwood	AF	75655	iu9@mail.com	233-212-2311	$1134	M	2/12/1987	10/6/2014
2015	ANH NJA	Vupx Avenue	Sunflower	FQ	72647	ag6@mail.com	918-192-1239	$3352	F	5/19/1998	6/14/2019
2016	AMJ JKI	VFUS Street	San Jose	QD	69021	ag4@rmail.com	712-432-1283	$2997	M	9/5/2000	9/21/2019
2017	MJN KOI	Nizz Drive	Santa Monica	FR	16384	xyz@email.com	817-927-2729	$4414	M	4/14/1981	8/7/2024

Input Patient Data

1. **Identifying columns that are ideal for mapping Appointments to Patients. Following provides best match %: CID, Email, CellPhone, Birthday, and Client Name.**
2. **This data is subjected to several data cleaning & transformations to comply with output data specifications.**

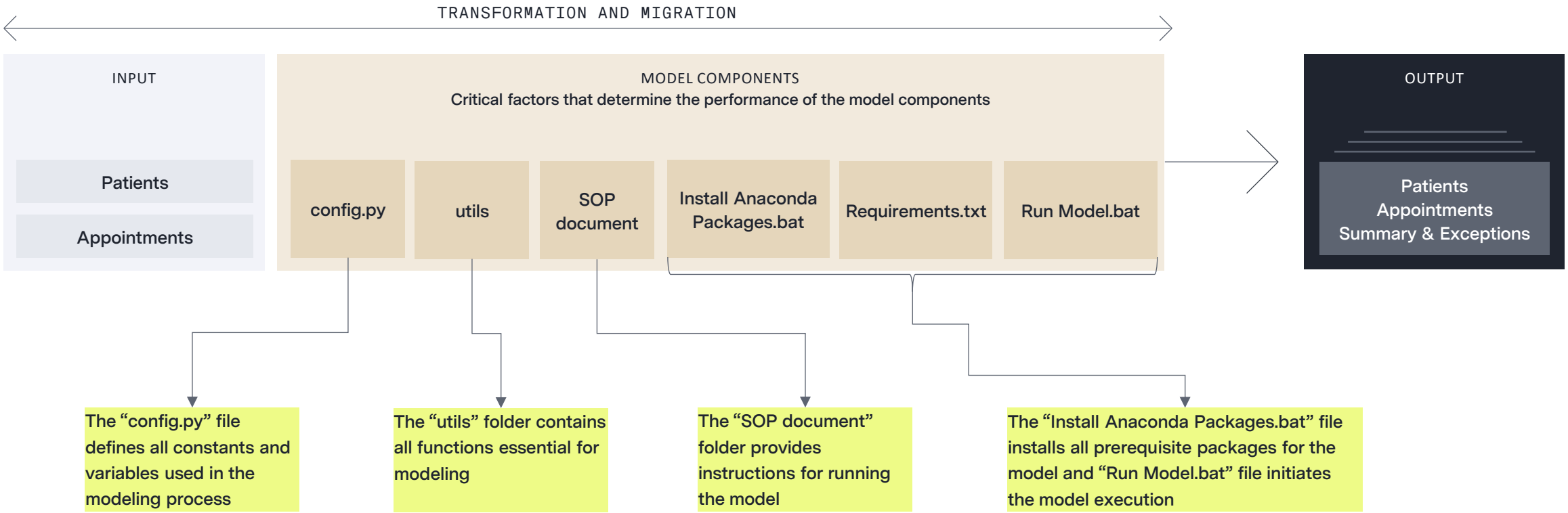
Input Appointment Data

1. **Identifying columns that are best for mapping Appointments to Patients. Following provides best match %: CID, Email, CellPhone, Birthday, and Name.**
2. **This data is subjected to several data cleaning & transformation steps to comply with output data specifications.**

CID	Name	Email	CellPhone	Gender	Birthday	Appt_Date	Appt_Status	Service_Name	StartTime	Staff_Name	Location
2011	ABC XYZ	abc@mail.com	234-734-2321	M	5/5/1988	6/22/2020	Cancelled	Meditation	12:42	Sophia	Test1
2012	BCD WXY	ag2@mail.com	917-232-2394	F	6/30/1997	2/1/2020	Invoiced	Botox	10:12	Mason	Test2
2013	ABH BHJ	ge2@mail.com	713-238-3231	M	6/27/1989	5/9/2020	Invoiced	DiamondG	7:11	Isabella	Test3
2014	ABH NJA	iu9@mail.com	233-212-2311	M	2/12/1987	11/23/2019	Invoiced	Check-up	3:17	Emma	Test4
2015	ANH NJA	ag6@mail.com	918-192-1239	F	5/19/1998	3/3/2020	No Show	Meditation	12:54	Isabella	Test7
2016	AMJ JKI	ag4@rmail.com	712-432-1283	M	9/5/2000	10/7/2020	Invoiced	Detox	4:36	Sophia	Test4
2017	MJN KOI	xyz@email.com	817-927-2729	M	4/14/1981	11/17/2020	Invoiced	Detox	18:14	Ethan	Test4
2014	ABH NJA	iu9@mail.com	234-734-2321	M	2/12/1987	9/22/2019	Invoiced	DiamondG	18:59	Mason	Test3
2014	ABH NJA	iu9@mail.com	917-232-2394	M	2/12/1987	9/26/2019	Invoiced	Check-up	15:13	Emma	Test2

Model snippet

The following framework aims to describe the model components and their relevance.



Output patients and appointments data

Patient ID	First Name	Last Name	Address 1	City	State	Zip	Birthdate	Gender	Cell Phone	Email	Provider Name
243	ABC	XYZ	Xytz Drive	Springdale	TX	33206	5/5/1988	M	234-734-2321	abc@mail.com	Jim
244	BCD	WXY	Qwib Lane	Pinecrest	TY	86651	6/30/1997	F	917-232-2394	ag2@mail.com	Jim
245	ABH	BHJ	Zorn Avenue	Willow	CA	45314	6/27/1989	M	713-238-3231	ge2@mail.com	Jim
246	ABH	NJA	Fyx Street	Oakwood	AF	75655	2/12/1987	M	233-212-2311	iu9@mail.com	Jim
247	ANH	NJA	Vupx Avenue	Sunflower	FQ	72647	5/19/1998	F	918-192-1239	ag6@mail.com	Jim
248	AMJ	JKI	VFUS Street	San Jose	QD	69021	9/5/2000	M	712-432-1283	ag4@rmail.com	Jim
249	MJN	KOI	Nizz Drive	Santa Monica	FR	16384	4/14/1981	M	817-927-2729	xyz@email.com	Jim

Output Patient Data

1. **Patient ID is a system-generated field with starting patient ID as user input.**
2. **Provider Name, which is a mandatory field is provided by the user as input., if not already available in Input dataset.**
3. **To avoid information loss, the CID from the input data is relocated to Custom Info 1.**

Output Appointment Data

1. **If a unique identification is found, the Patient ID is generated after matching with the patient's data.**
2. **All the non-mapped appointments are added to the exception file.**
3. **The values in the ApptStatus in the input file are used to generate the entries for Is Cancelled, Is Confirmed, and Is No Show fields.**

Patient ID	Appointment Date	Provider	Start Time	Duration	Is Cancelled	Location	Reason	Is Confirmed	Is No Show
243	6/22/2020	Sophia	12:42	6	FALSE	Test1	Meditation	TRUE	NULL
244	2/1/2020	Mason	10:12	25	TRUE	Test2	Botox	TRUE	NULL
245	5/9/2020	Isabella	7:11	12	FALSE	Test3	DiamondG	TRUE	TRUE
246	11/23/2019	Emma	3:17	38	NULL	Test4	Check-up	NULL	TRUE
247	3/3/2020	Isabella	12:54	45	NULL	Test7	Meditation	TRUE	NULL
248	10/7/2020	Sophia	4:36	14	NULL	Test4	Detox	NULL	NULL
249	11/17/2020	Ethan	18:14	11	TRUE	Test4	Detox	NULL	NULL
243	9/22/2019	Mason	18:59	41	FALSE	Test3	DiamondG	TRUE	NULL
245	9/26/2019	Emma	15:13	11	NULL	Test2	Check-up	TRUE	NULL

Summary & exceptions file

Notes

1. This sheet contains the appointment summary statistics, matching breakdown of appointments with patients and the breakdown of appointments and patients from the input to output transition.

2. The cells highlighted in "grey" are the number of records that have been removed from the input appointments and patients datasets.

Breakdown of Appointments	
Total Input Appointments	20
Duplicate Appointments	0
Appointment End Time Crossed EOD	0
Appointments with Empty Required Fields	0
Total Unmatched Appointments	0
Total Output Appointments	20
Invalid Field Format Appointments	5

Matching Breakdown	
Total Output Appointments (A)	20
Total Office Appointments (B)	0
Total Unmatched Appointments (C)	0
Total Matched Appointments (D)	20
Match Percentage (%) ((D/(A-B+C))*100)	100

Breakdown of Patients	
Total Input Patients	20
Duplicate Patients	0
Patients With Empty Required Fields	0
Archived/Inactive Patients	0
Total Output Patients	20
Invalid Field Format Patients	7

Appointments Data Summary

1. Matching Breakdown gives information about the match percentage between appointments and patients, as well as the non-patient appointments.
2. Patients and Appointments are further broken down and summarized based on type of data exception.

Patient Field Exceptions

1. The invalid/inaccurate data is removed from the cell, but the row is retained.
2. Exception Description explains why the cell is blanked out.
3. The Is Purged column indicates if the row has been eliminated due to an exception, based on whether it's an exception in Mandatory or Non-Mandatory columns.

Index	CID	ClientName	Address	City	State	Postal	Email	HomePhone	CellPhone	Balance	Gender	Birthday	Exception Description	Is Purged
0	2011	ABC XYZ	Xylz Drive	Springdale	TX	38098	abc@mail.com	234-734-2321	234-734-2321	4960	M	1986-02-14	Invalid CellPhone Format	No
0	2011	ABC XYZ	Xylz Drive	Springdale	TX	38098	abc@mail.com	917-232-2394	917-232-2394	4960	M	1986-02-14	Invalid HomePhone Format	No
3	2014	DEF UVW	Fyx Street	Oakwood	TX	65146	def@mail.com	713-238-3231	713-238-3231	-40	M	1989-11-30	Invalid CellPhone Format	No
7	2018	HIJ QRS	Tylp Lane	Sunflower	TX	13677	hij@gmail.com	233-212-2311	233-212-2311	-55	M	1993-12-02	Invalid CellPhone Format	No
9	2020	JKL OPQ	Zyrk Drive	Briarwood	TX	51694	jkl@mail.com	918-192-1239	918-192-1239	-400	F	1995-04-25	Invalid CellPhone Format	No
11	2022	LMN MNO	Quyz Avenue	W Grove	TX	17221	ln@hmail.com	712-432-1283	712-432-1283	349	M	1997-10-05	Invalid CellPhone Format	No
12	2023	MNO LMN	Xyzk Street	Lakeside	TX	73569	oln@hmail.com	817-927-2729	817-927-2729	54	F	1998-02-28	Invalid HomePhone Format	No
14	2025	OPQ JKL	Blaz Lane	P Valley	TX	90573	opl@hmail.com	234-734-2321	234-734-2321	109	F	2000-08-04	Invalid HomePhone Format	No

Notes

1. This sheet contains all the patients with format issues in the field values.

2. The "Index" column indicates the row number of the record in raw data. The error is there in the "Exception Description" Column against each record.

Learnings

1) **Acquired knowledge of the healthcare Industry:**
    - Gained in-depth understanding of Electronic Health Record (EHR) systems and the types of data they manage
    - Enhanced industry expertise through HIPAA certification, ensuring compliance with healthcare data privacy and security standards

2) **Enhanced working knowledge of Python**:
    - Gained deeper insights into the features and functionalities of packages like pandas, numpy , tkinter , openpyxl and xlsxwriter.
    - Improved proficiency in using Anaconda, Jupyter notebook and workflows for data migration and transformation.
    - Learnt how to develop the code for creating batch files to automate a series of commands.
    - Learnt how to decode regular expressions and use them in validating data fields based on their output requirements.
    - Learnt how to automate creation of excel files from python code as per the output requirements.