github.com/JackEdwardAuty/Markdown-Notes-Processing-and-Analysis/

Markdown Processing Personal Project

Intro – Data Collection & Tools

- Built in **Python**, utilising **DataFrame**, **Pandas** libraries
- **Markdown**: industry standard lightweight markup language used for documenting code bases, processes and readme files.
- **Obsidian Vault:** Custom markdown-based PKM system
  - templated daily entries for medication details, and wake hours

Personal-Knowledge-Management

- Custom File-Handling and *Data **Extract**ion* Pipeline
  - Custom Regex to handle varying data entry formats across Templates

```
48 extract_wake_times(process_file_headers(*get_file_info(test_case_full)),
     get_file_info(test_case_full))
49
```

```
[{'Sleep': '~',
  'Wake': '14:45',
  'Time': '',
  'Quality': 'Decent',
  'date': '2025-11-30',
  'filename': '2025-11-30.md',
  'folder': '2025-11',
  'section_type': 'Wake+Sleep'}]
```

Wake

Table of Contents

Sleep

Sleep	Wake	Time	Quality
~	{{time}}		Decent

```
Processing 1513 markdown files...
Saved wake_time.csv with 118 rows
Saved medication.csv with 901 rows
```

```
53 extract_med_times(process_file_headers(*get_file_info(test_case_full)),
     get_file_info(test_case_full))
```

```
[{'Time': '17:09',
  'Medication': 'Ritalin IR',
  'Dosage': '5mg',
  'date': '2025-11-30',
  'filename': '2025-11-30.md',
  'folder': '2025-11',
  'section_type': 'Prescription'},
 {'Time': '17:09',
  'Medication': 'Fexofenadine',
```

Medication

Table of Contents

Prescribed

Time	Medication	Dosage
{{time}}	Ritalin XR	2x18mg
	Ritalin IR	5mg

github.com/JackEdwardAuty/Markdown-Notes-Processing-and-Analysis/

Markdown Processing Personal Project

- Custom *Data* ***Extraction*** & ***Normalization*** Pipeline
  - **Process:** Additional bespoke regex patterns for Normalisation parse_dosage_string
  - Pandas transformations in Python; output standardized format for Tableau.

- ***Transforming*** the extracted data to ease further analysis
  - **Non-Standard Formats Handled:**
    - **Dosage:**
      - 2x18mg → 36mg, [Ritalin XR]
      - 1 (tab) → 500mg, 1.5x500mg → 750mg [Paracetamol]
    - **Multiple Times:**
      - 15:25, 19:40 → separate rows (same dosage)
    - **Partial Doses:**
      - 17:15, 21:35 (half) → separate rows (adjust dosage)
    - **Time Ranges:**
      - 09:45-50 → 09:47 (midpoint)
    - **Non-Standard:**
      - logged to notes column for manual review

To implement these Transformations, I introduce three new methods:

- `parse_dosage_string(dosage_str, factor=1.0)`
  - Seven Regex patterns to handle varied input form
- `parse_time_value(time_str)`
  - Handles time ranges and non-standard to notes column
- `split_times_to_rows(dataframe)`
  - Handles splitting across multiple rows with multiplicative factors

```
def parse_dosage(dosage_str, factor=1.0):
    """
    Handles:
      - 2x18mg, 1.5x5mg, 1.5x500mg
      - 5mg, 120mg, 500mg
      - 2 x tab, 2 tab, 2 tabs
      - 1 (tab), 2 (tabs)
    Returns (amount, unit) or (None, original) if no match.
    """
```

- CSV now files ready for **Load** stage
  - Tableau for Visualisation and further Analysis

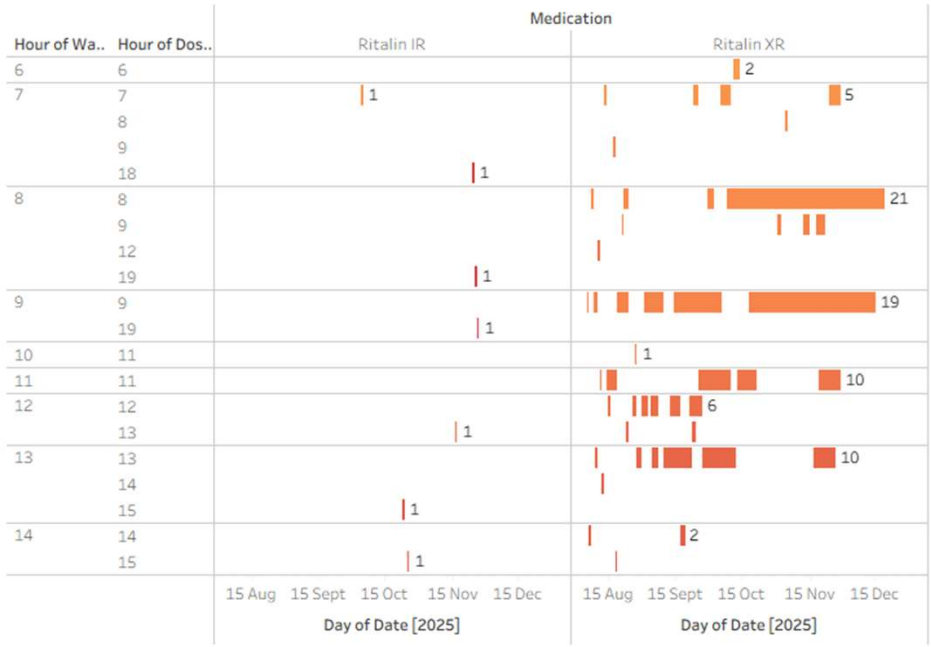
Markdown Processing Tableau Dashboard

Medication Timing, Wake Patterns, and Adherance Streaks

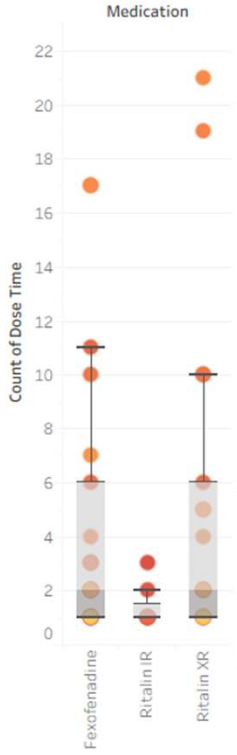
github.com/JackEdwardAuty/Markdown-Notes-Processing-and-Analysis/

Exploring how my Wake Time and Dose Timing relate to Ritalin Adherance

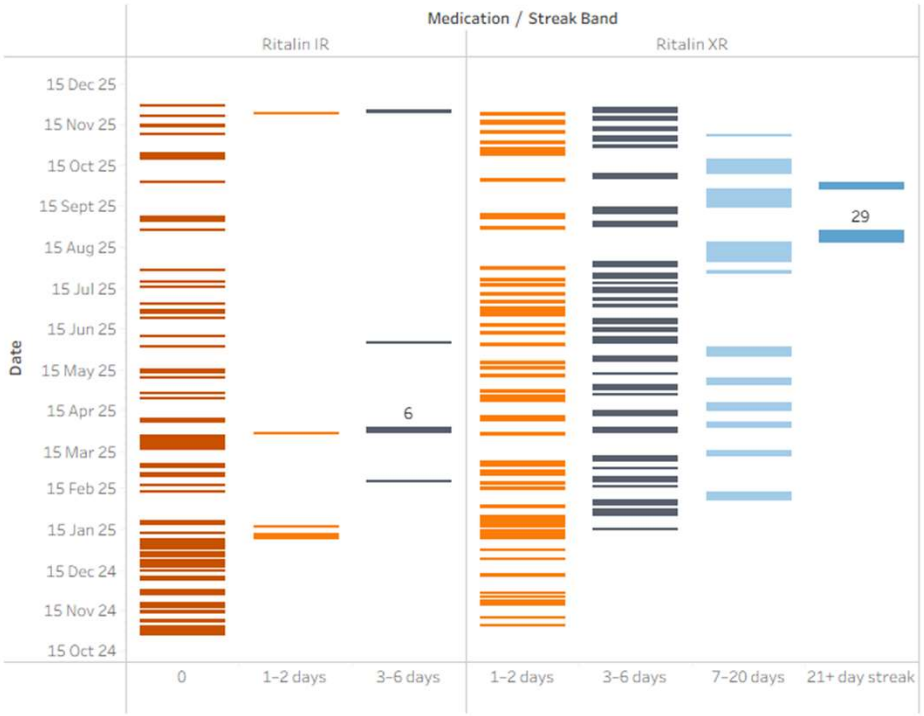
Dose Time vs Wake Time (Prescription)



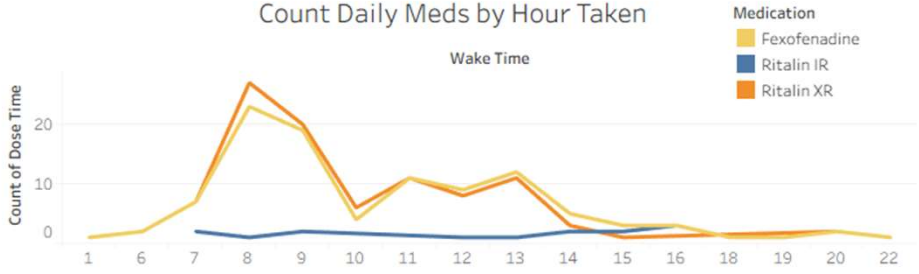
Variation in Dose Hour



Streak Counter by Streak Band (Prescription)



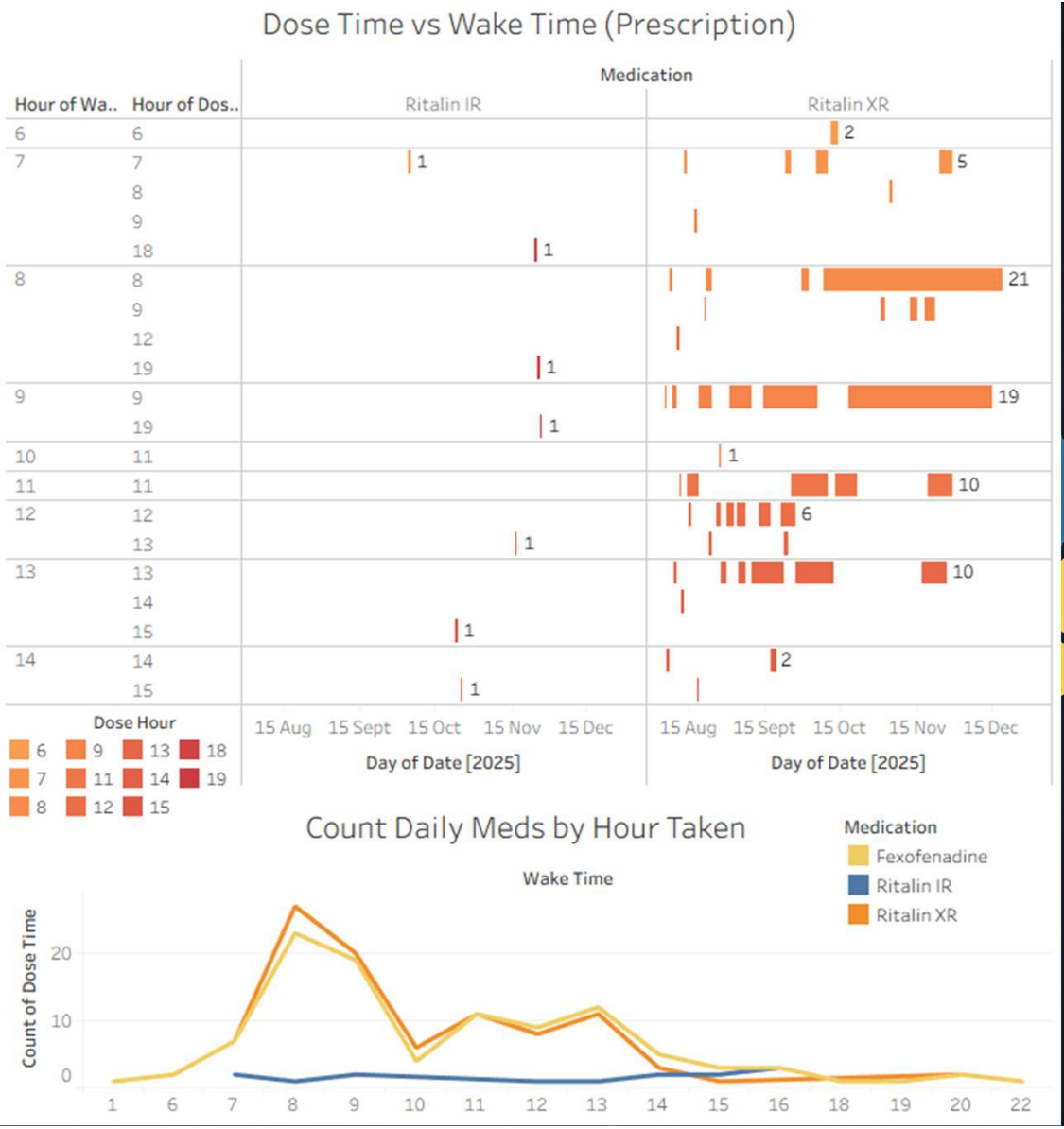
Count Daily Meds by Hour Taken



Medication Streak by Wake Hour

Medication	Streak Band	Wake Time 1	6	7	8	9	10	11	12	13	14	15	20	22
Ritalin XR	1–2 days			5	7	6	2		1	4		1		
	3–6 days			3	46	37	11	3	9	9			9	
	7–20 days		15	47	97	88	7	108	43	49	35			11
	21+ day streak	25			66	21	51	72	51	54				

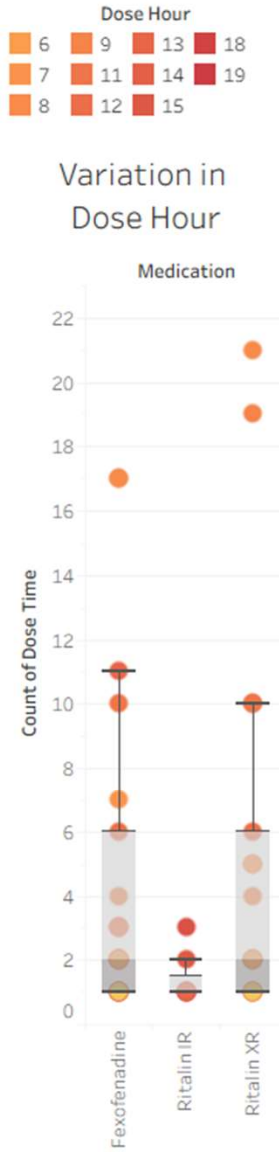
Streak Band
0
1–2 days
3–6 days
7–20 days
21+ day streak

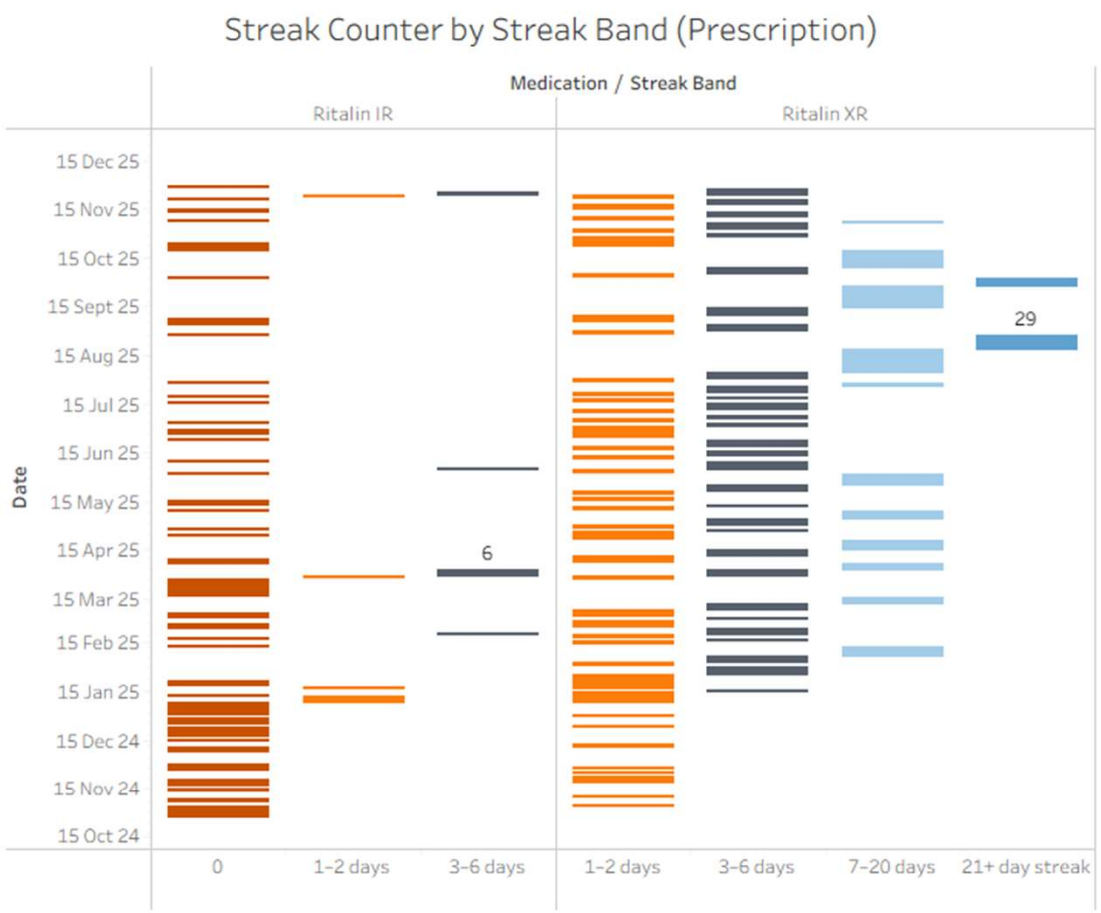


Dashboard Takeaways

- Dose Time against Wake Time Using RUNNING_SUM method
  - Found that most consistently took XR medication at 8 AM and 9 AM
  - More likely to take (late-day) Ritalin IR the earlier I woke up
- Count Daily Meds by Hour Taken
  - Stayed consistent with Fexofenadine whether woke up too late for Ritalin XR
  - Confirmation of 8-9 AM consistency

github.com/JackEdwardAuty/Markdown-Notes-Processing-and-Analysis/





Medication Streak by Wake Hour

Medication	Streak Band	Wake Time 1	6	7	8	9	10	11	12	13	14	15	20	22
Ritalin XR	1–2 days			5	7	6	2		1	4		1		
	3–6 days			3	46	37	11	3	9	9			9	
	7–20 days		15	47	97	88	7	108	43	49	35			11
	21+ day streak	25			66	21	51	72	51	54				

Streak Band
0
1–2 days
3–6 days
7–20 days
21+ day streak

Dashboard Takeaways

- Variation in Dose Hour
  - Box-and-Whisker showing distribution
- Streak Counter by Band
  - Longest XR medication streak was 29 days in a row
  - Ritalin IR was rarely taken multiple days in a row
- Med Streak by Wake Hour
  - Longest runs (most consistent distribution) occurs when I wake between 8-10am

github.com/JackEdwardAuty/Markdown-Notes-Processing-and-Analysis/