

Conquer Experience Presentation

Shang-shiun Tsai

12/1/2023

Agenda

- ❖ Introduction
- ❖ Key findings
- ❖ How we approach
- ❖ Dataset overview
- ❖ Business questions we solved
- ❖ Business insights/Future recommendation

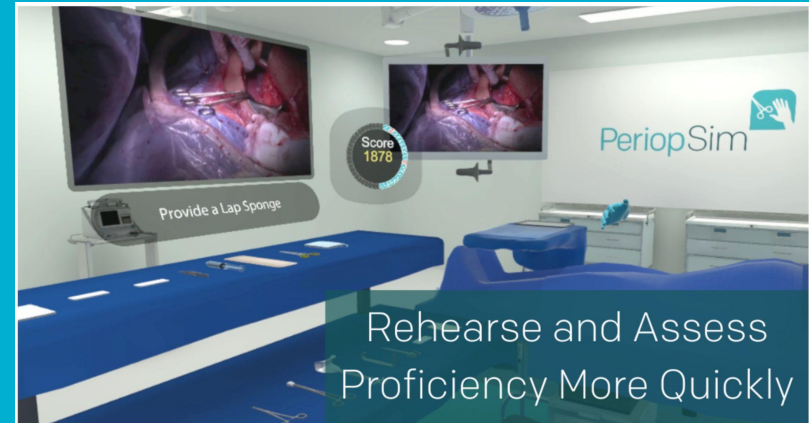
Introduction

Conquer Experience created a virtual reality tool named PeriopSim about practicing surgery procedures. It makes nurses and surgical technologists productive more quickly by practicing clinical skills before they enter operation room.

Our project aims to gain value insights into learner preferences, engagement patterns, and user experiences by using learner's data analysis.

Sponsor's Questions Overview:

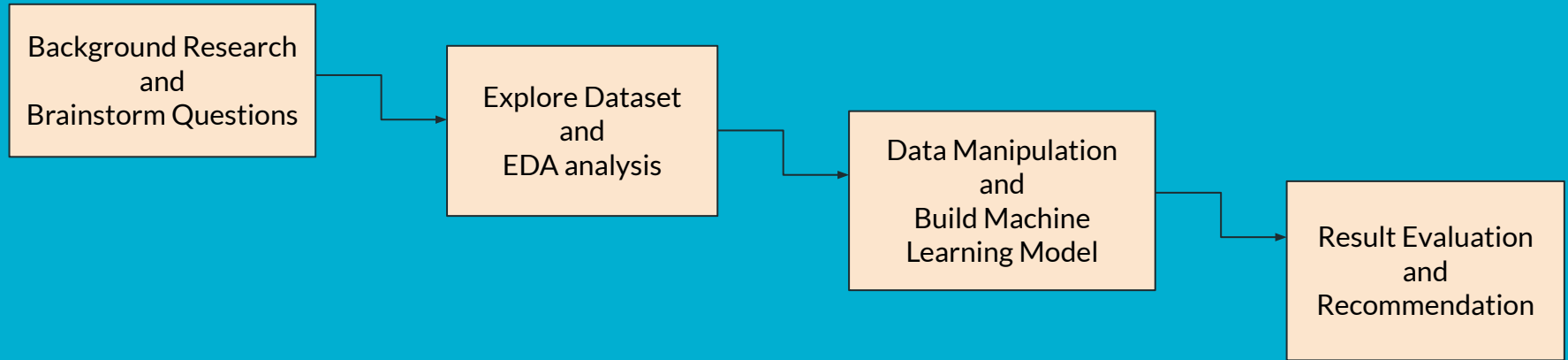
1. How the VR tools help educators to improve performance?
2. How to distinguish good and bad learners?
3. What are the key characteristics will affect performance?
4. Time prediction of achieving specific proficiency performance



Key Findings

1. Users performance improved when they repeatedly use our tools multiple times.
2. Design two standard to evaluate the best and worst users.
3. Conducted features engineering to create 6 variables to fit in machine learning models in order to predict the performance (CorrectRate). Achieved the R-square of 41.4% for the Linear Regression Model and RMSE (Error) of 0.079
4. Developed a method to estimate the Number of Practice Times Needed to Achieve 10-90% Proficiency.



How we approach



Tools: Python, Tableau

Method: Correlation Analysis, Linear Regression Model

Data Overview – Data Schema

<i>First level</i>		<i>Second Level</i>		<i>Third Level</i>
Row_id	one to many	Range_id	many to one	Summary_Correct_try
Scenario_id		Tools_id		Summary_Error_try
User_id		Tools_selected		Summary_Timesaved
DateTime		Metrics: (Correct_First_try Error_first_try Timesaved)		

Scenario: Bipolar Hip 1/4 Instrumentation, Hip Nail 3/3 Procedure, Burr Hole Procedure

Tasks: 'Provide a Stapler', 'Irrigation please', 'Scissors please'

Data Overview: Convert Json to xlsx

```
"total_rows": 1289944,
"offset": 743840,
"rows": [{
  "id": "log:63b0ce0b-b95f2b0c-3d08",
  "key": ["account:k7do45g3-5mghf0jt", 1672531468507],
  "value": null,
  "doc": {
    "_id": "log:63b0ce0b-b95f2b0c-3d08",
    "_rev": "1-dfd4d338e7813f256515cb8d38d1d196",
    "id": "63b0ce0b-b95f2b0c-3d08",
    "context": {
      "scenario": "bundle:5aa2b6e0-7b0a8f61",
      "scenario_name": "Hip Nail 3/3 Instrumentation",
      "started": 1672531171623,
      "finished": 1672531467960,
      "user": "user:lb5ht6fs-kl3ik40h",
      "realm": "account:k7do45g3-5mghf0jt",
      "environment": {
        "app_type": "Desktop",
        "app_name": "PeriopSim",
        "app_version": "1.15.3",
        "platform": "iPhonePlayer",
        "os": "iPadOS 16.1",
        "app_settings": {
          "randomized_instruments": true
        }
      }
    }
  }
}]
```



rows.id	rows.doc.context.scenario_name	rows.doc.data.summary.time_saved	SessionStartDateTime	Session_day
log:64025c5a-cf4b9e33-56bf	Hip Nail 3/3 Procedure	10.500000	2023-03-23 18:06:04	2023-03-23
log:64025c5a-cf4b9e33-56bf	Hip Nail 3/3 Procedure	10.500000	2023-03-23 18:06:04	2023-03-23
log:64025c5a-cf4b9e33-56bf	Hip Nail 3/3 Procedure	10.500000	2023-03-23 18:06:04	2023-03-23
log:64025c5a-cf4b9e33-56bf	Hip Nail 3/3 Procedure	10.500000	2023-03-23 18:06:04	2023-03-23
log:64025c5a-cf4b9e33-56bf	Hip Nail 3/3 Procedure	10.500000	2023-03-23 18:06:04	2023-03-23
...
log:64ee697f-9a8db5af-3455	Inguinal Hernia Repair Procedure	49.711889	2023-09-19 22:30:53	2023-09-19
log:64ee697f-9a8db5af-3455	Inguinal Hernia Repair Procedure	49.711889	2023-09-19 22:30:53	2023-09-19
log:64ee697f-9a8db5af-3455	Inguinal Hernia Repair Procedure	49.711889	2023-09-19 22:30:53	2023-09-19
log:64ee697f-9a8db5af-3455	Inguinal Hernia Repair Procedure	49.711889	2023-09-19 22:30:53	2023-09-19
log:64ee697f-9a8db5af-3455	Inguinal Hernia Repair Procedure	49.711889	2023-09-19 22:30:53	2023-09-19

140478 rows x 76 columns

rows.doc.data.summary.time_saved	rows.doc.data.summary.time_lost	rows.doc.data.summary.correct_first_try	rows.doc.data.summary.errors_count
10.500000	0.000000	26.0	0.0
10.500000	0.000000	26.0	0.0
10.500000	0.000000	26.0	0.0
10.500000	0.000000	26.0	0.0
10.500000	0.000000	26.0	0.0
...
49.711889	5.394669	49.0	0.0
49.711889	5.394669	49.0	0.0
49.711889	5.394669	49.0	0.0
49.711889	5.394669	49.0	0.0
49.711889	5.394669	49.0	0.0

Single User Dataset Overview

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 139534 entries, 0 to 140477
Data columns (total 63 columns):
```

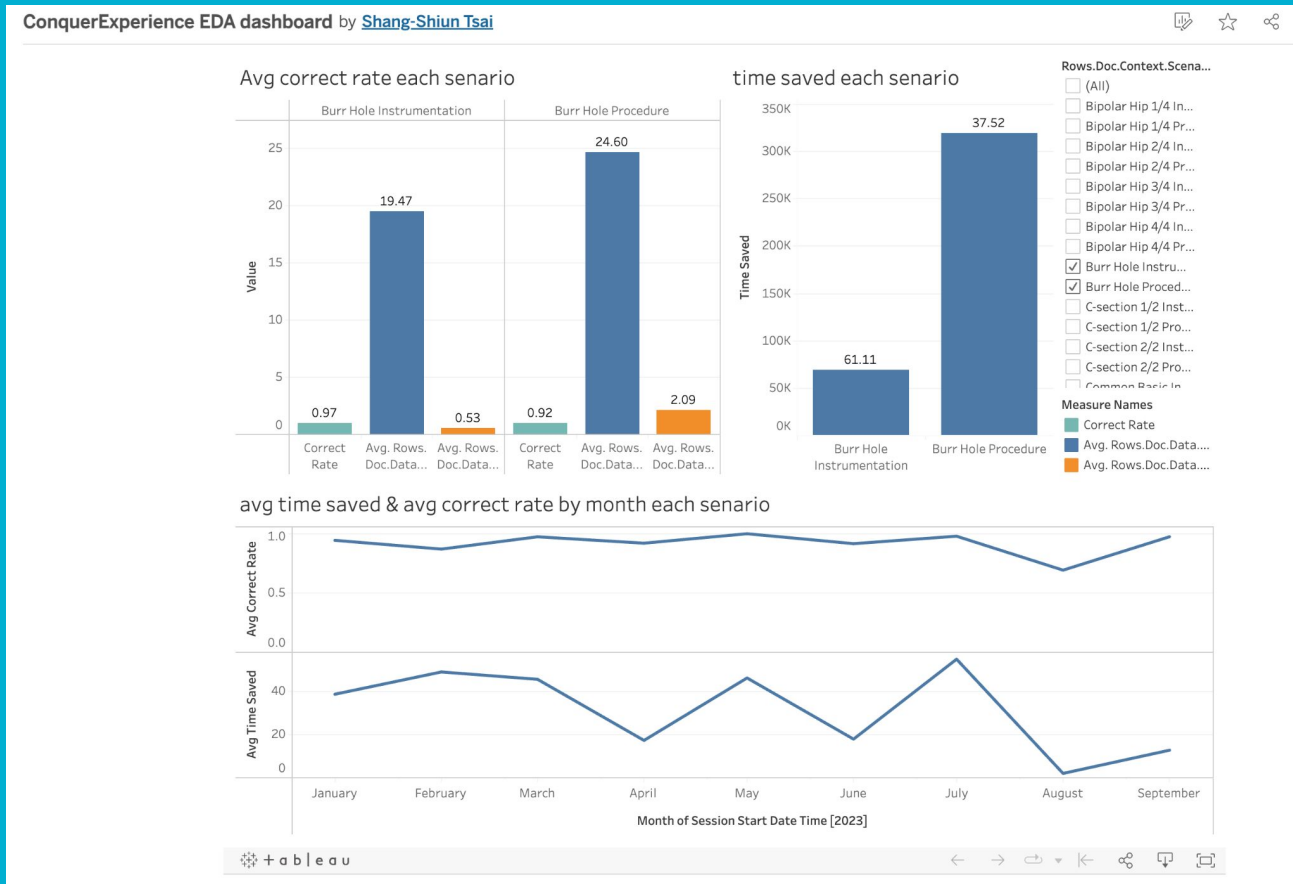
#	Column	Non-Null Count	Dtype
0	rows.id	139534 non-null	object
1	rows.key.0	139534 non-null	object
2	rows.key.1	139534 non-null	int64
3	rows.doc._id	139534 non-null	object
4	rows.doc._rev	139534 non-null	object
5	rows.doc.id	139534 non-null	object
6	rows.doc.context.scenario	139534 non-null	object
7	rows.doc.context.scenario_name	139534 non-null	object
8	rows.doc.context.started	139534 non-null	int64
9	rows.doc.context.finished	139534 non-null	int64
10	rows.doc.context.user	139534 non-null	object
11	rows.doc.context.realm	139534 non-null	object
12	rows.doc.context.environment.app_type	139534 non-null	object
13	rows.doc.context.environment.app_name	139534 non-null	object
14	rows.doc.context.environment.app_version	139534 non-null	object
15	rows.doc.context.environment.platform	139534 non-null	object
16	rows.doc.context.environment.os	139534 non-null	object
17	rows.doc.context.environment.app_settings.randomized_instruments	139534 non-null	bool
18	rows.doc.context.environment.login_mode	139534 non-null	object
19	rows.doc.context.environment.hw.configuration	139534 non-null	object
20	rows.doc.context.environment.hw.gpu	139534 non-null	object
21	rows.doc.context.environment.hw.xr_refresh_rate	139534 non-null	int64
22	rows.doc.context.environment.hw.xr_tracking_space	139534 non-null	object
23	rows.doc.context.environment.hw.processor	139534 non-null	object
24	rows.doc.data.events.range_id	139346 non-null	object
25	rows.doc.data.events.range_start_time	139346 non-null	float64
26	rows.doc.data.events.instructions	139346 non-null	object
27	rows.doc.data.events.tool_ids	139346 non-null	object
28	rows.doc.data.events.metrics.expected_duration	139346 non-null	float64
29	rows.doc.data.events.metrics.duration	139346 non-null	float64
30	rows.doc.data.events.metrics.time_lost	139346 non-null	float64
31	rows.doc.data.events.metrics.time_saved	139346 non-null	float64

32	rows.doc.data.events.metrics.tools_selected	139346 non-null	object
33	rows.doc.data.events.metrics.instrument_correct_first_try	139346 non-null	float64
34	rows.doc.data.events.metrics.instrument_errors	139346 non-null	float64
35	rows.doc.data.events.metrics.score	139346 non-null	float64
36	rows.doc.data.events.metrics.passing.destination	139346 non-null	object
37	rows.doc.data.events.metrics.passing.expectedDestination	139346 non-null	object
38	rows.doc.data.events.metrics.passing.velocity_magnitude	139346 non-null	float64
39	rows.doc.data.events.metrics.passing.velocity.x	139346 non-null	float64
40	rows.doc.data.events.metrics.passing.velocity.y	139346 non-null	float64
41	rows.doc.data.events.metrics.passing.velocity.z	139346 non-null	float64
42	rows.doc.data.events.metrics.passing.angle_delta.x	139346 non-null	float64
43	rows.doc.data.events.metrics.passing.angle_delta.y	139346 non-null	float64
44	rows.doc.data.events.metrics.passing.angle_delta.z	139346 non-null	float64
45	rows.doc.data.events.metrics.passing.translation_delta.x	139346 non-null	float64
46	rows.doc.data.events.metrics.passing.translation_delta.y	139346 non-null	float64
47	rows.doc.data.events.metrics.passing.translation_delta.z	139346 non-null	float64
48	rows.doc.data.events.type	139534 non-null	object
49	rows.doc.data.events.timestamp	139534 non-null	float64
50	rows.doc.data.events.score	139534 non-null	float64
51	rows.doc.data.summary.track_id	139534 non-null	object
52	rows.doc.data.summary.instrument_count	139534 non-null	float64
53	rows.doc.data.summary.score	139534 non-null	float64
54	rows.doc.data.summary.duration	139534 non-null	float64
55	rows.doc.data.summary.time_saved	139534 non-null	float64
56	rows.doc.data.summary.time_lost	139534 non-null	float64
57	rows.doc.data.summary.correct_first_try	139534 non-null	float64
58	rows.doc.data.summary.errors_count	139534 non-null	float64
59	rows.doc.data.type	139534 non-null	object
60	rows.doc.type	139534 non-null	object
61	rows.doc.created	139534 non-null	int64
62	SessionStartDateTime	139534 non-null	datetime64[ns]

dtypes: bool(1), datetime64[ns](1), float64(27), int64(5), object(29)
memory usage: 67.2+ MB

EDA dashboard for the dataset with 1 example user

Create new variable:
Correct Rate



Burr Hole Dataset Overview

	rows.doc_id	rows.doc.context.realm	rows.doc.context.scenario_name	rows.doc.context.scenario	rows.doc.context.user	rows.doc.data.events.instruct	ist	rows.doc.data.type	rows.doc.data.summary.time_saved	total_rows	SessionStartDateTime	CorrectRate	TimeSaved_std	SessionStartYear	SessionStartMonth	
	0	log:645c217c-21d48b91-d9e5	account:00000000-00000000	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:00055993-Medcerts	Needle Driver, 4-0 Vicryl Rapic Adson For	00	playthrough	61.826124	1309139	2023-06-02 02:40:00	1.000000	2.377928	2023	2023-06-01
	1	log:645d5584-1f8e2445-3c3f	account:00000000-00000000	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:00053815-Medcerts	Needle Driver, 4-0 Vicryl Rapic Adson For	00	playthrough	78.997627	1309139	2023-06-01 08:00:00	1.000000	3.038370	2023	2023-06-01
	2	log:645d6e72-e1d1931e-023a	account:00000000-00000000	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:00055548-Medcerts	Needle Driver, 4-0 Vicryl Rapic Adson For	00	playthrough	74.677595	1309139	2023-06-01 08:00:00	1.000000	2.872215	2023	2023-06-01
	3	log:645f9c9-c23d36da-bf7a	account:00000000-00000000	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:00052340-Medcerts	Needle Driver, 4-0 Vicryl Rapic Adson For	00	playthrough	17.757771	1309139	2023-06-02 13:20:00	1.000000	0.682991	2023	2023-06-01
	4	log:646018ef-db51bf8e-8a3a	account:00000000-00000000	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:00055812-Medcerts	Needle Driver, 4-0 Vicryl Rapic Adson For	00	playthrough	8.045076	1309139	2023-06-05 08:00:00	1.000000	0.309426	2023	2023-06-01
	
1048570	log:622e3e2a-0bf851cd-aa54	account:kcm93knk-q9aIf2rx	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:ktx80zwc-89svmmvw	Needle Driver, 3-0 Vicryl & Jeffe For	72	playthrough	66.038445	1309139	2022-04-01 13:20:00	0.961538	2.539940	2022	2022-04-01	
1048571	log:5d5215d0-b9344104-b71c	account:jwtzuas8-9we8s7e2	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:j234yt1s-o5d4ke0r	Jefferson Forceps and Monoq Cal	51	playthrough	81.065910	1309139	2019-08-14 13:20:00	0.961538	3.117920	2019	2019-08-01	
1048572	log:5d5215d0-b9344104-b71c	account:jwtzuas8-9we8s7e2	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:j234yt1s-o5d4ke0r	Jefferson Forceps and Monoq Cal	51	playthrough	81.065910	1309139	2019-08-14 13:20:00	0.961538	3.117920	2019	2019-08-01	
1048573	log:5d5215d0-b9344104-b71c	account:jwtzuas8-9we8s7e2	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:j234yt1s-o5d4ke0r	Jefferson Forceps and Monoq Cal	51	playthrough	81.065910	1309139	2019-08-14 13:20:00	0.961538	3.117920	2019	2019-08-01	
1048574	log:5d5215d0-b9344104-b71c	account:jwtzuas8-9we8s7e2	Burr Hole Procedure	bundle:5388ee79-0c4f9590	user:j234yt1s-o5d4ke0r	Jefferson Forceps and Monoq Cal	51	playthrough	81.065910	1309139	2019-08-14 13:20:00	0.961538	3.117920	2019	2019-08-01	
1048575 rows x 31 columns																

Burr Hole Dataset Overview

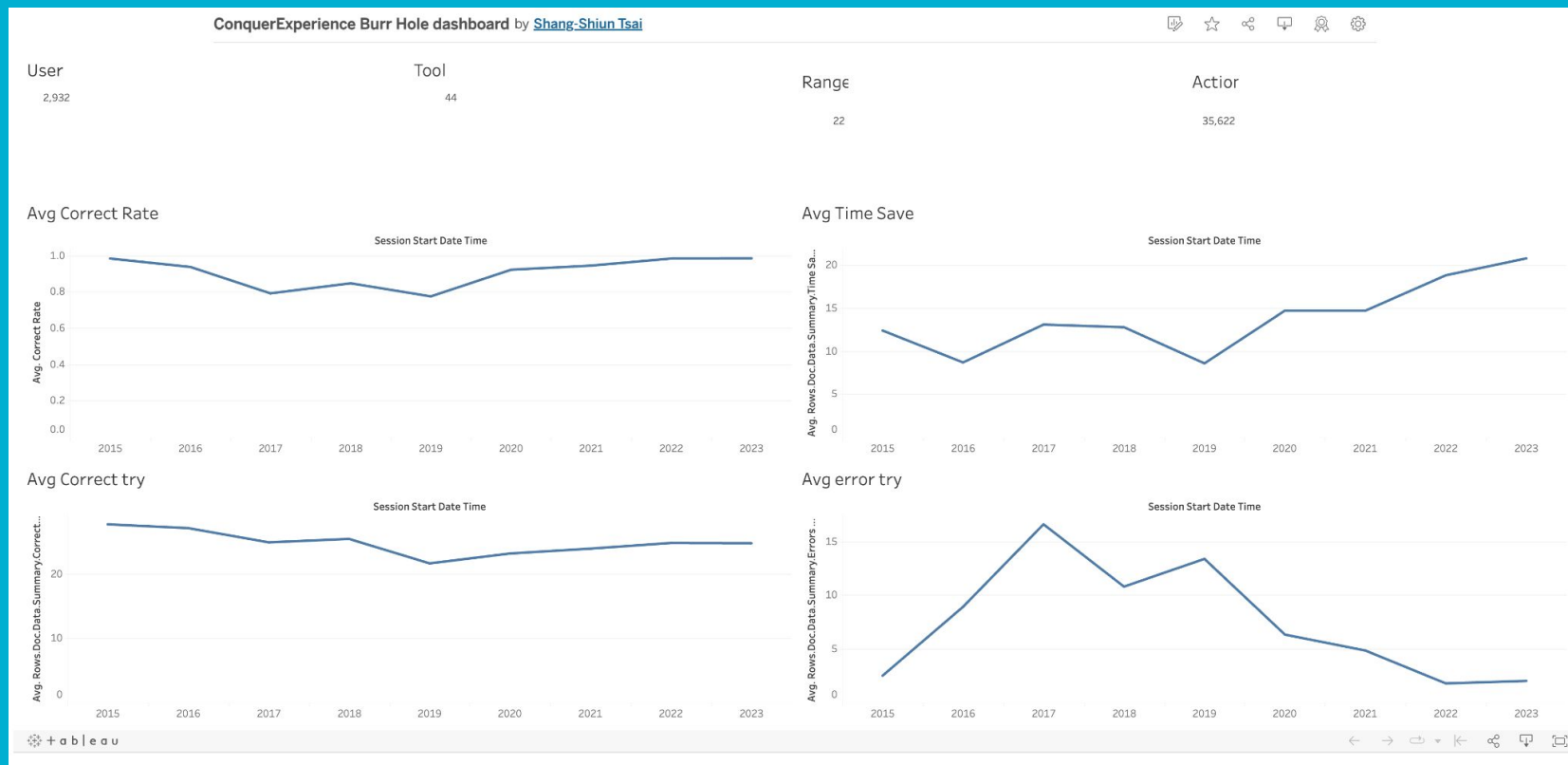
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1048575 entries, 0 to 1048574  
Data columns (total 31 columns):
```

#	Column	Non-Null Count	Dtype
0	rows.doc._id	1048575 non-null	object
1	rows.doc.context.realm	1048575 non-null	object
2	rows.doc.context.scenario_name	1048575 non-null	object
3	rows.doc.context.scenario	1048575 non-null	object
4	rows.doc.context.user	1047463 non-null	object
5	rows.doc.data.events.instructions	1040199 non-null	object
6	rows.doc.data.events.metrics.duration	1040199 non-null	float64
7	rows.doc.data.events.metrics.expected_duration	1040199 non-null	float64
8	rows.doc.data.events.metrics.instrument_correct_first_try	1040199 non-null	float64
9	rows.doc.data.events.metrics.instrument_errors	1040199 non-null	float64
10	rows.doc.data.events.metrics.score	1040199 non-null	float64
11	rows.doc.data.events.metrics.time_lost	1040199 non-null	float64
12	rows.doc.data.events.metrics.time_saved	1040199 non-null	float64
13	rows.doc.data.events.metrics.tools_selected	1040199 non-null	object
14	rows.doc.data.events.range_id	1040199 non-null	object
15	rows.doc.data.events.tool_ids	1040199 non-null	object
16	rows.doc.data.events.type	1040958 non-null	object
17	rows.doc.data.summary.correct_first_try	1040962 non-null	float64
18	rows.doc.data.summary.duration	1040962 non-null	float64
19	rows.doc.data.summary.errors_count	1040962 non-null	float64
20	rows.doc.data.summary.instrument_count	1040962 non-null	float64
21	rows.doc.data.summary.score	1040962 non-null	float64
22	rows.doc.data.summary.time_lost	1040962 non-null	float64
23	rows.doc.data.type	1048575 non-null	object
24	rows.doc.data.summary.time_saved	1040962 non-null	float64
25	total_rows	1048575 non-null	int64
26	SessionStartDateTime	1048575 non-null	datetime64[ns]
27	CorrectRate	1040958 non-null	float64
28	TimeSaved_std	1040962 non-null	float64
29	SessionStartYear	1048575 non-null	int64
30	SessionStartMonth	1048575 non-null	datetime64[ns]

dtypes: datetime64[ns](2), float64(16), int64(2), object(11)
memory usage: 248.0+ MB

EDA dashboard for the Burr Hole dataset



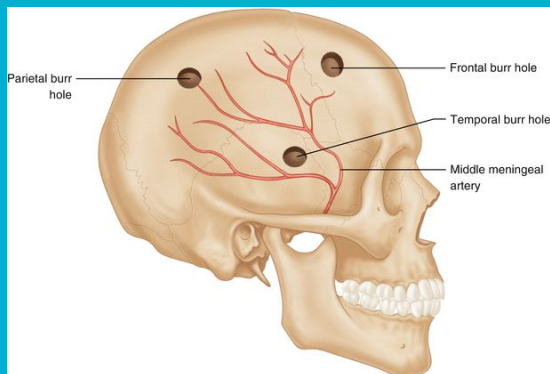
Define Best/Worst User in Burr Hole Dataset

Method 1:

Best: CorrectRate > 0.9, worst: CorrectRate < 0.5

Highest CorrectRate value: 1.0

Lowest CorrectRate value: 0.00757575757575758



Distinct users with CorrectRate > 0.9:

```
rows.doc.context.user
user:00000000-0000011a    90
user:00000000-0000012c   154
user:00000000-0000021a    53
user:00000000-0000022a    52
user:00000000-0000022c   140
```

```
...
user:lmhsjgkf-e9n1d9e4   163
user:lmih4o7h-85x4yl5u    86
user:lmjk6y40-oh8q1vwn    94
user:lp3m9b0-x2sblrzg     76
user:lp3nsjc-zqsmxclh     82
```

Name: CorrectRate, Length: 2633, dtype: int64

Distinct users with CorrectRate < 0.5:

```
rows.doc.context.user
user:00000000-0001062c    88
user:00000000-00aaaa1a    81
user:00000005-nquer QA   214
user:00018985-Medcerts   150
user:00038921-Medcerts   134
```

```
...
user:ljhhjlgj-gcfw9zex   132
user:ll5ga4hs-x5qy4b09   238
user:ll5gdmitt-eyaraca   278
user:lmmaxfa0w-t9qrgrbg   172
user:lmhsm2az-ri3q7457   124
```

Name: CorrectRate, Length: 152, dtype: int64

Define Best/Worst User in Burr Hole Dataset

Method 2:

Statistical Analysis, q1(0.25) and q3(0.75)

```
print(mean)
```

```
print(std_dev)
```

```
0.9246847478211807
```

```
0.1518033166887376
```

```
Threshold for best customers: 1.0764880645099182  
Threshold for worst customers: 0.7728814311324431
```

Best Customers:

	rows.doc.context.user	count
0	user:00000000-0000011a	90
1	user:00000000-0000012c	154
2	user:00000000-0000021a	53
3	user:00000000-0000022a	52
4	user:00000000-0000022c	140
...
2548	user:lm0s4cpt-d6avs0v9	697
2549	user:lmhsjgkf-e9n1d9e4	163
2550	user:lmih4o7h-85x4yl5u	86
2551	user:lm3m9b0-x2sblrzg	76
2552	user:lm3nsjc-zqsmxclh	82

```
[2553 rows x 2 columns]
```

Worst Customers:

	rows.doc.context.user	count
0	user:00000000-0000112c	66
1	user:00000000-0000122c	58
2	user:00000000-0000262c	61
3	user:00000000-0000382c	64
4	user:00000000-0000782c	62
...
534	user:llb6x64o-ygc42q8x	210
535	user:lldr3cnk-0ohufu9i	106
536	user:llm959va-2wmuhz20	126
537	user:lmmaxfa0w-t9qrgbrg	172
538	user:lmhsm2az-ri3q7457	124

```
[539 rows x 2 columns]
```


Data cleaning – Missing Value

Burr Hole Dataset: Before data cleaning

```
RangeIndex: 1048575 entries, 0 to 1048574  
Data columns (total 31 columns):
```

Remove about 8000 rows of duplicate, null value, and select 12 necessary columns

```
<class 'pandas.core.frame.DataFrame'>  
Int64Index: 1040199 entries, 0 to 1040198  
Data columns (total 12 columns):  
#   Column                                     Non-Null Count  Dtype  
---  ----  
0   rows.doc._id                             1040199 non-null object  
1   rows.doc.context.scenario_name           1040199 non-null object  
2   rows.doc.context.scenario                1040199 non-null object  
3   rows.doc.context.user                    1040199 non-null object  
4   rows.doc.data.summary.correct_first_try  1040199 non-null float64  
5   rows.doc.data.summary.errors_count       1040199 non-null float64  
6   rows.doc.data.summary.time_saved         1040199 non-null float64  
7   SessionStartDateTime                     1040199 non-null datetime64[ns]  
8   CorrectRate                             1040199 non-null float64  
9   TimeSaved_std                           1040199 non-null float64  
10  SessionStartYear                         1040199 non-null int64  
11  SessionStartMonth                       1040199 non-null datetime64[ns]  
dtypes: datetime64[ns](2), float64(5), int64(1), object(4)  
memory usage: 103.2+ MB
```


Data Cleaning – Outliers

Figure 1: Distribution of CorrectRate

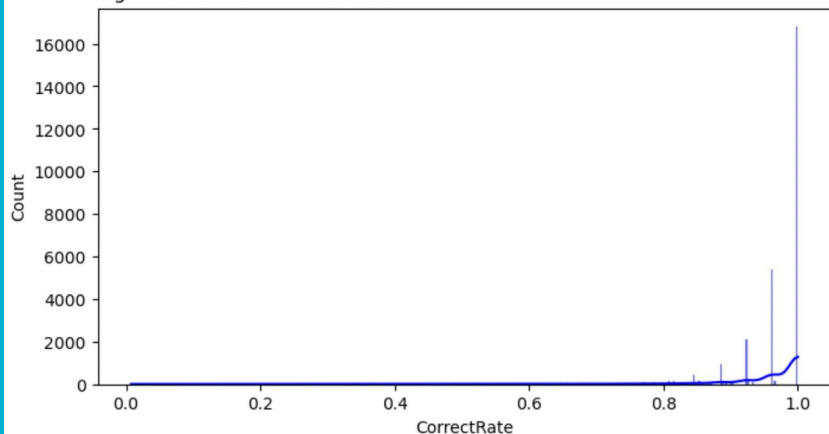


Figure 2: Distribution of CorrectRate_adj_p1

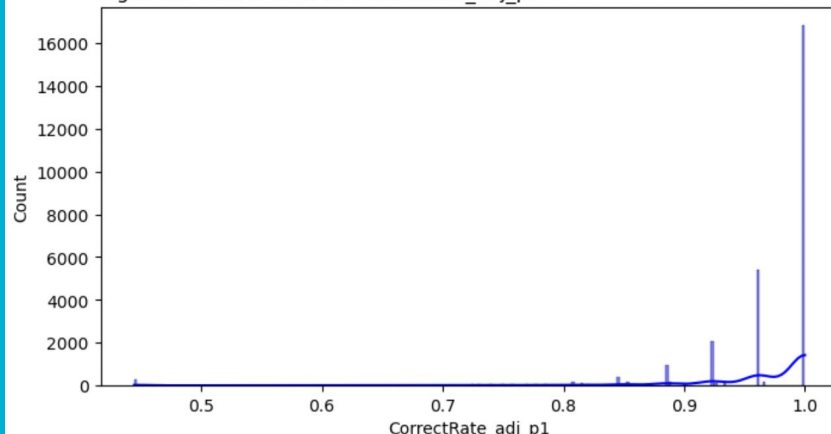
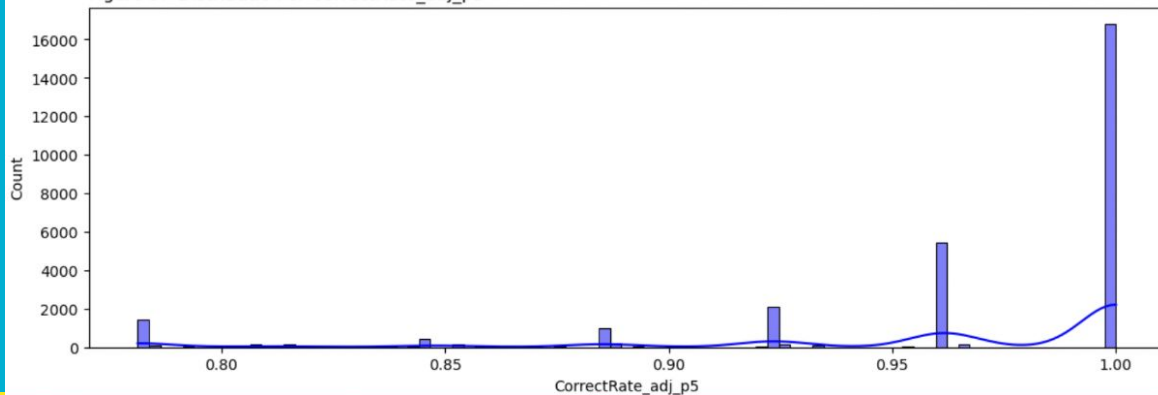


Figure 3: Distribution of CorrectRate_adj_p5



Data Cleaning – Data Overlap

Proper Data

rows.doc.data.events.tool_ids	rows.doc.data.events.metrics.tools_selected
[tool:5a9ebc85-be2820ac, tool:5a9ebf57-ceb6220b,tool:5a9ec264-a2a4282d]	[tool:5a9ebc85-be2820ac, tool:5a9ebf57-ceb6220b,tool:5a9ec264-a2a4282d]

Actual Data

rows.doc.data.events.tool_ids	rows.doc.data.events.metrics.tools_selected
tool:5a9ebc85-be2820ac	tool:5a9ebc85-be2820ac
tool:5a9ebc85-be2820ac	tool:5a9ebf57-ceb6220b
tool:5a9ebc85-be2820ac	tool:5a9ec264-a2a4282d
tool:5a9ebf57-ceb6220b	tool:5a9ebc85-be2820ac
tool:5a9ebf57-ceb6220b	tool:5a9ebf57-ceb6220b
tool:5a9ebf57-ceb6220b	tool:5a9ec264-a2a4282d
tool:5a9ec264-a2a4282d	tool:5a9ebc85-be2820ac
tool:5a9ec264-a2a4282d	tool:5a9ebf57-ceb6220b
tool:5a9ec264-a2a4282d	tool:5a9ec264-a2a4282d

Machine Learning – Features Definition

CorrectRate = Summary Correct First Try / (Summary Correct First Try + Summary Error First Try)

Timesaved_std: Time save of each session

Average_Previous_CorrectRate: Average of all previous session of that User

Previous_CorrectRate: The most recent Correct Rate of that User

Recency: The time between now and the last session

Repetition Time: Number of practice times before the current session

Number of tools in Session: Number of tools in the session

Number of Range in Session: Number of range/task in the session

```
model_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
Index: 28550 entries, 353 to 10655
```

```
Data columns (total 16 columns):
```

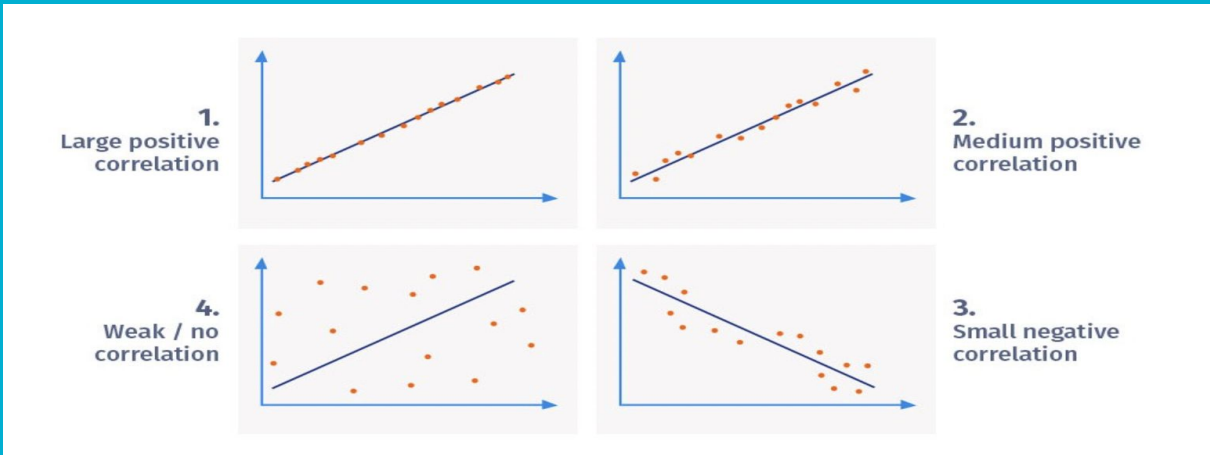
#	Column	Non-Null Count	Dtype
0	TimeSaved_std	28550 non-null	float64
1	average_prev_timesaved	28550 non-null	float64
2	prev_TimeSaved_std	28550 non-null	float64
3	CorrectRate	28550 non-null	float64
4	average_prev_correctrate	28550 non-null	float64
5	prev_CorrectRate	28550 non-null	float64
6	CorrectRate_adj_p1	28550 non-null	float64
7	average_prev_correctrate_adj_p1	28550 non-null	float64
8	prev_CorrectRate_adj_p1	28550 non-null	float64
9	CorrectRate_adj_p5	28550 non-null	float64
10	average_prev_correctrate_adj_p5	28550 non-null	float64
11	prev_CorrectRate_adj_p5	28550 non-null	float64
12	recency	28550 non-null	float64
13	repetition_time	28550 non-null	int64
14	number_of_range_in_session	28550 non-null	int64
15	number_of_tools_in_session	28550 non-null	int64

```
dtypes: float64(13), int64(3)
```

```
memory usage: 3.7 MB
```

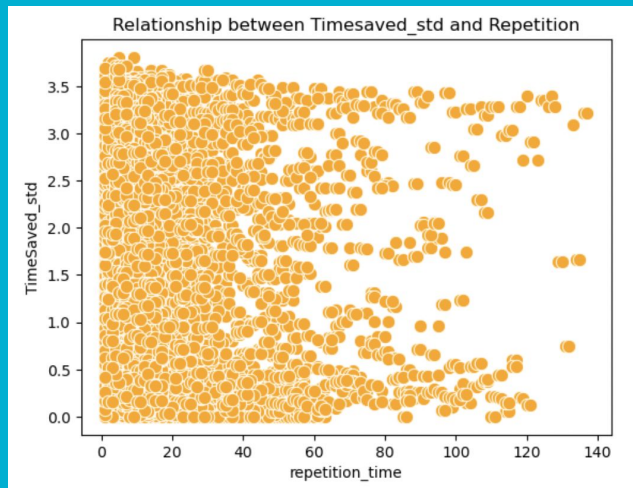
Correlation Analysis

Correlation analysis assesses the strength and direction of a linear relationship between two or more variables.

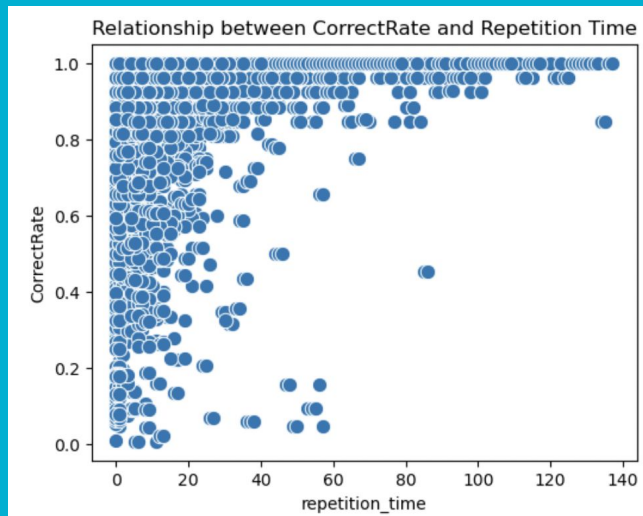


Correlation Analysis

	CorrectRate
CorrectRate	1.000000
average_prev_correctrate	0.054738
prev_CorrectRate	0.173408
recency	-0.074622
repetition_time	0.057562
number_of_tools_in_session	-0.098092
number_of_range_in_session	-0.100217



0.144



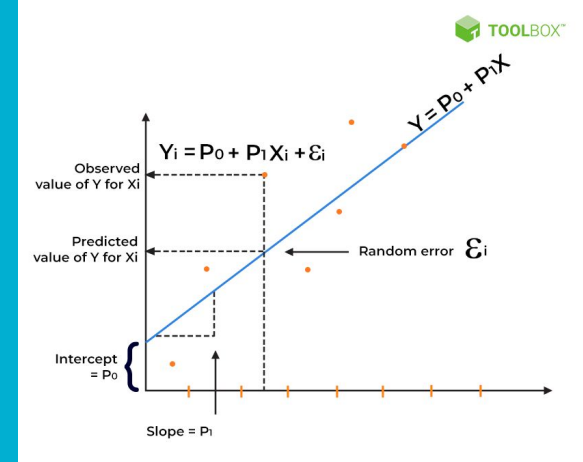
Machine Learning – Linear Regression

What is linear regression?

Linear regression models the relationship between variables, revealing how changes in one predict changes in another, useful for predictions and insights.

Purpose : Helps identify the impact of factors on business outcomes, aiding in predictive analysis and decision-making.

CorrectRate = **x1** * Average_Previous_CorrectRate + **x2***
Previous_CorrectRate - **x3** * Recency + **x4***
Repetition_Time + **x5***Number_of_Tools_in_Session - **x6***
Number_of_Range_in_Session



Machine Learning Linear – Linear Regression

Linear Regression - All Data

$\text{CorrectRate} = 0.9535 + 0.0013 * \text{Average_Prev_CorrectRate}$
 $+ 0.0165 * \text{Previous_CorrectRate} - 0.0079 * \text{Recency} + 0.0019$
 $* \text{Repetition_Time} + 0.0271 * \text{Number_of_Tools_in_Session} -$
 $0.0347 * \text{Number_of_Range_in_Session}$

1. How model fit : $R_{\text{square}} = 4.3\%$

2. Accuracy of prediction model:
 $\text{Test_RMSE} = 6.532$

Linear Regression - Repetition > 0

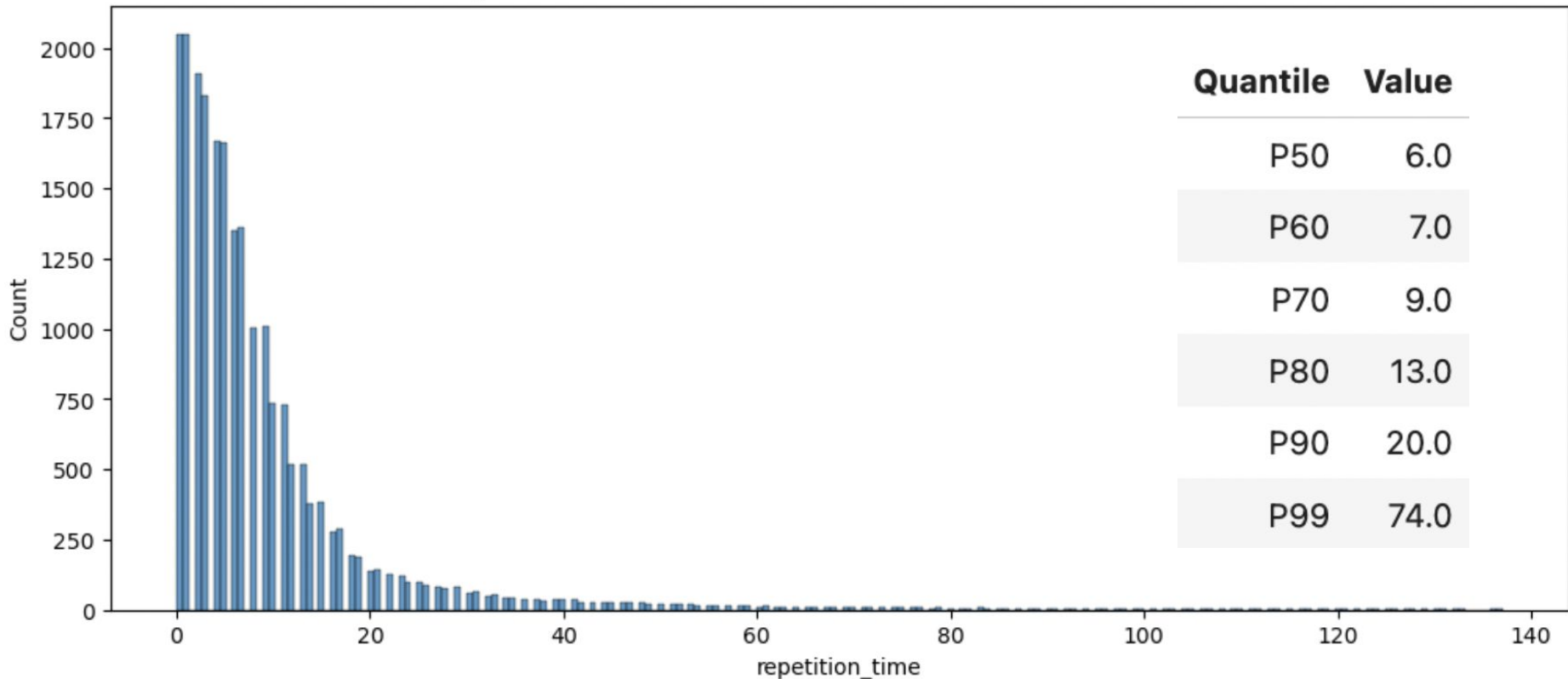
$\text{CorrectRate} = 0.9568 + 0.0020 * \text{Average_Prev_CorrectRate}$
 $+ 0.0622 * \text{Previous_CorrectRate} - 0.0024 * \text{Recency} + 0.0026$
 $* \text{Repetition_Time} + 0.0229 * \text{Number_of_Tools_in_Session} -$
 $0.0241 * \text{Number_of_Range_in_Session}$

1. How model fit : $R_{\text{square}} = 40.4\%$

2. Accuracy of prediction model:
 $\text{Test_RMSE} = 0.079$

Predict Number of Practice Times Needed to Achieve 95% Proficiency

Repetition Time Distribution of User with CorrectRate > 95%



Business insights

1. Positive impact: Previous correct rate, repetition time and number of tools.
Negative impact: Recency and number of range in session.
2. Prediction of the number of simulations provides insights for optimizing learning paths and personalized interventions in training programs.
3. Efficiency improves, saving time during surgical practices.
4. Improved surgical skills and decreased error rates.

Future recommendations

1. Compare with “leaderboard” users
2. Conduct cross-platform performance analysis
3. Explore top-performing users' patterns
4. Enhance predictions with demographics
5. Develop complicated ML models

Thank you