

JANUARY 10TH, 2021

# Data Processing Challenge

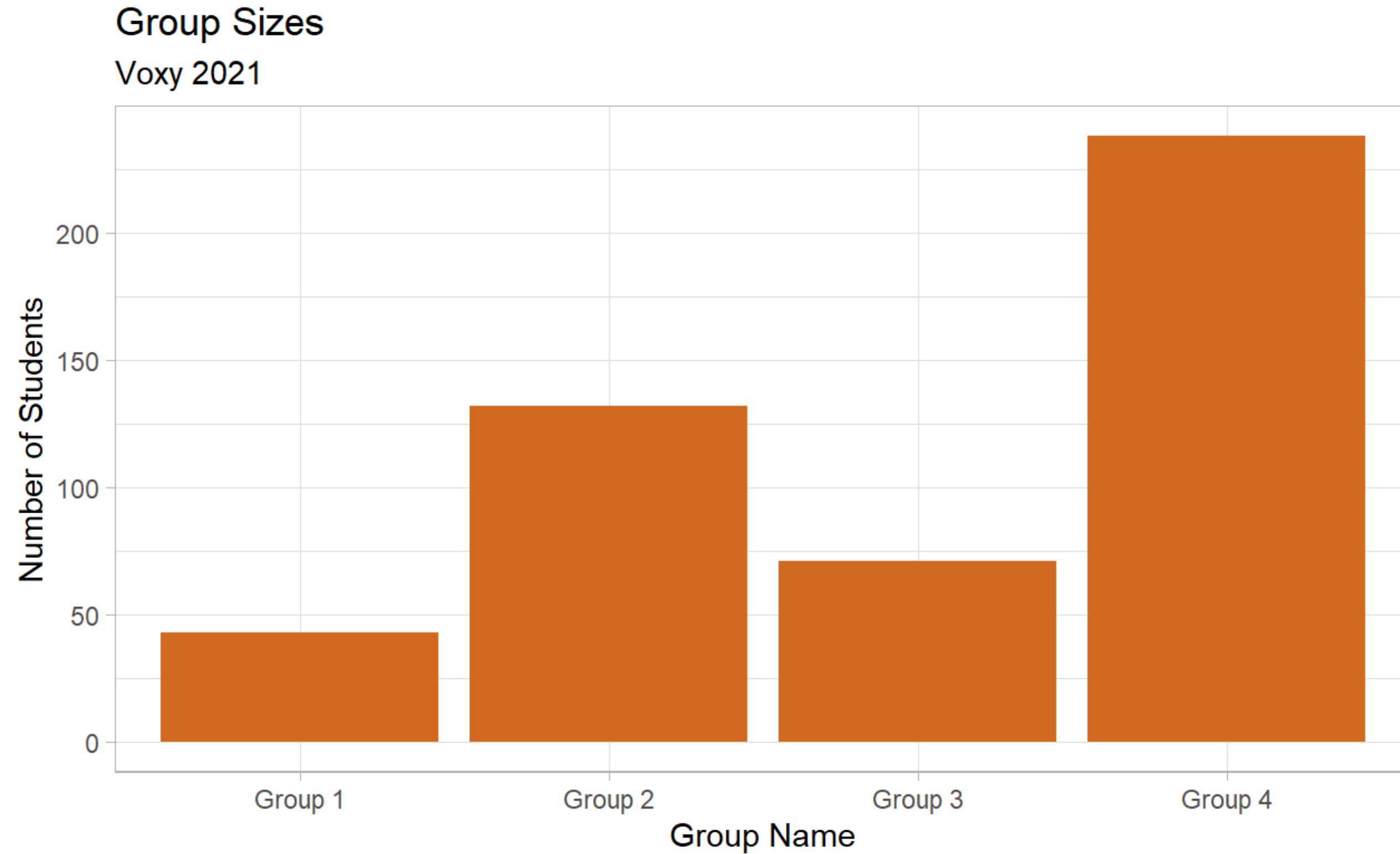
Voxy

# Data

- Voxy students
- October - December (2021)
- 485 rows
- 22 columns

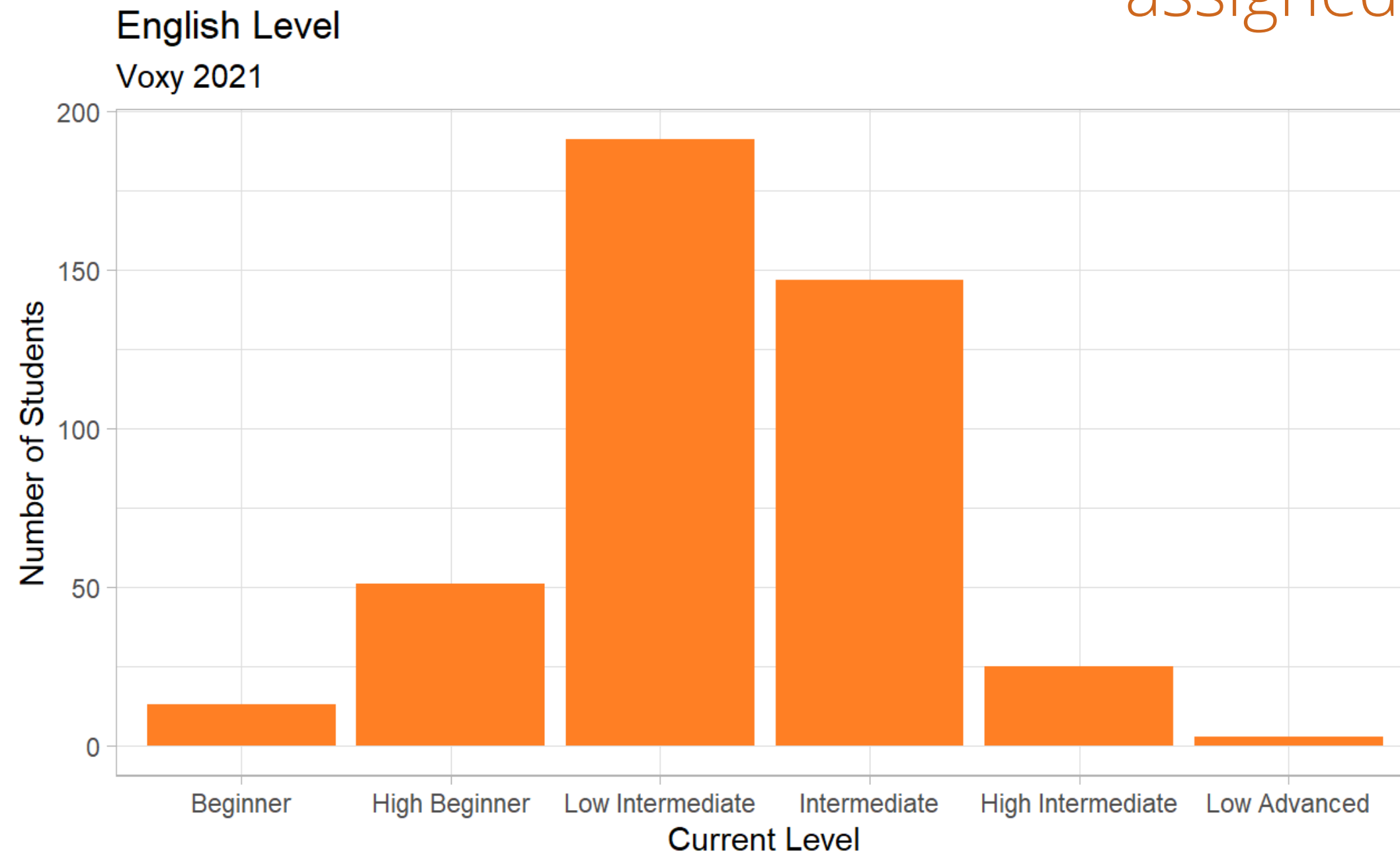
	A	B	C	D
1	User	group_name	current_level_name	last_login_date
2	User 1	Group 4		
3	User 10	Group 3	Low Intermediate	2021-12-05
4	User 100	Group 4	Beginner	2021-12-13
5	User 101	Group 4	Intermediate	2021-12-07
6	User 102	Group 4	High Intermediate	2021-12-13
7	User 103	Group 4	Intermediate	2021-12-15
8	User 104	Group 4	Intermediate	2021-12-17
9	User 105	Group 4	High Intermediate	2021-12-20
10	User 106	Group 4		
11	User 107	Group 4	High Intermediate	2021-11-22
12	User 108	Group 4	High Beginner	2021-12-02

All students in the dataset have  
1-year licenses with Voxy



The group sizes are different!

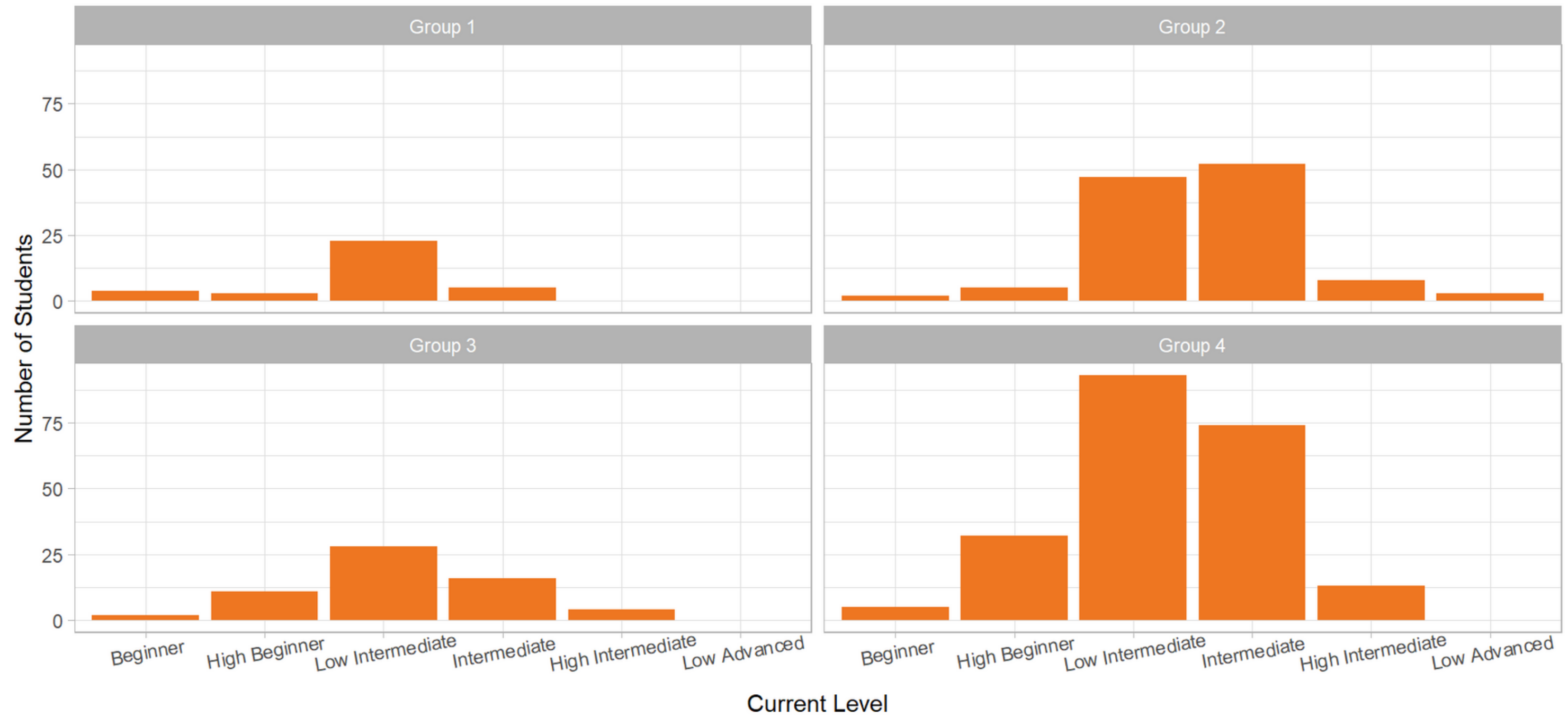
There are 54 students not assigned to any level.



Most of the students are in intermediate levels.

## Level Distribution per Group

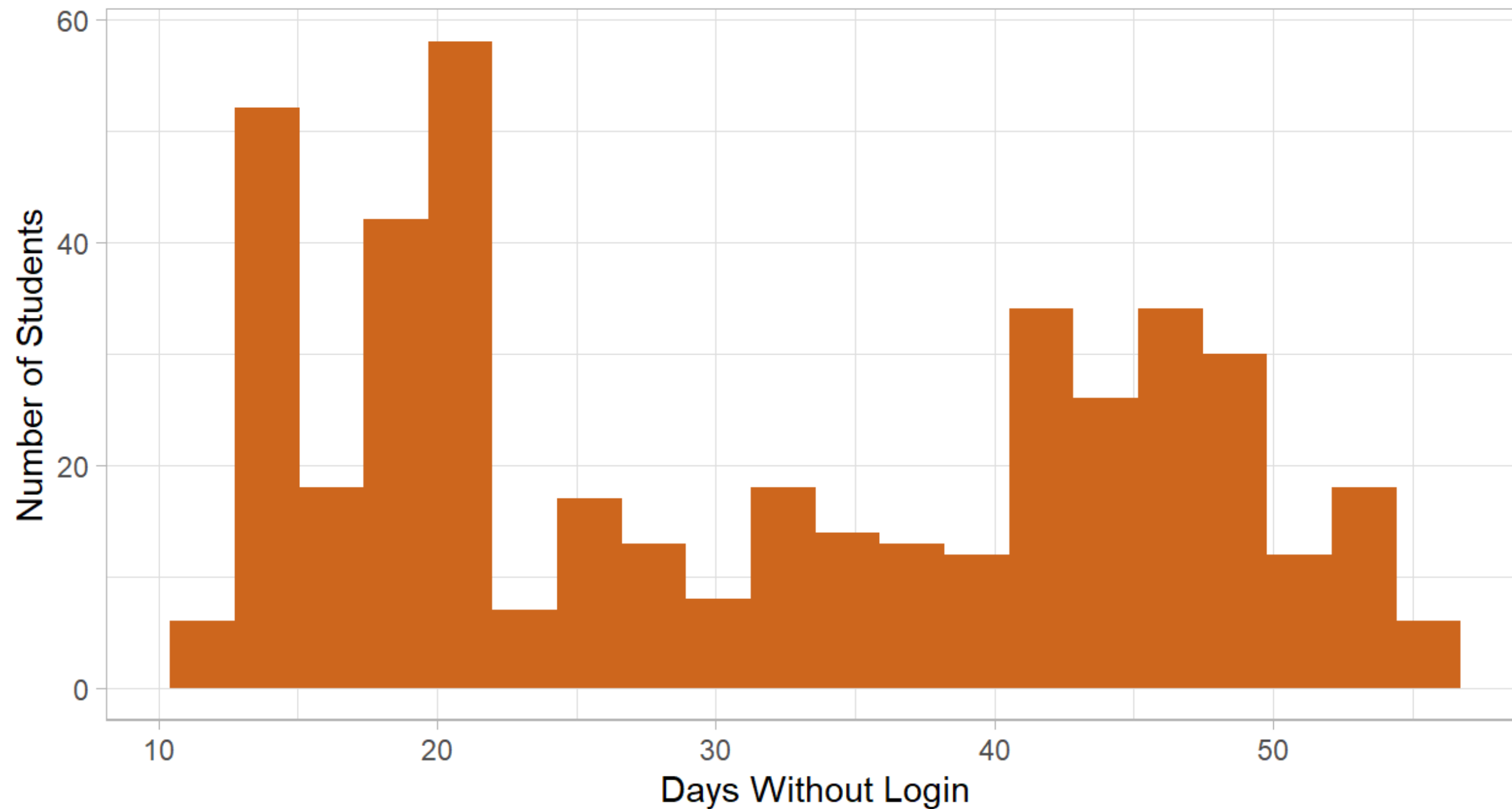
Voxy 2021



Level distribution is similar between groups.

## Distribution of Days Without Login

Voxy 2021



Days without login = January 4th - last login date.

Many students spent less than 20 days without logging in. But there were also quite a few that stayed between 40 and 50 days logged out.

# Summary Statistics for Days Without Login

Mean	Min	1 <sup>st</sup> Quartile	Median	3 <sup>rd</sup> Quartile	Max
31.5 days	12 days	19 days	30.5 days	43.8 days	56 days

Students spend, on average, 31 and a half days without logging in.



Which  
variables  
affect the  
completion  
of  
activities?

Study time;

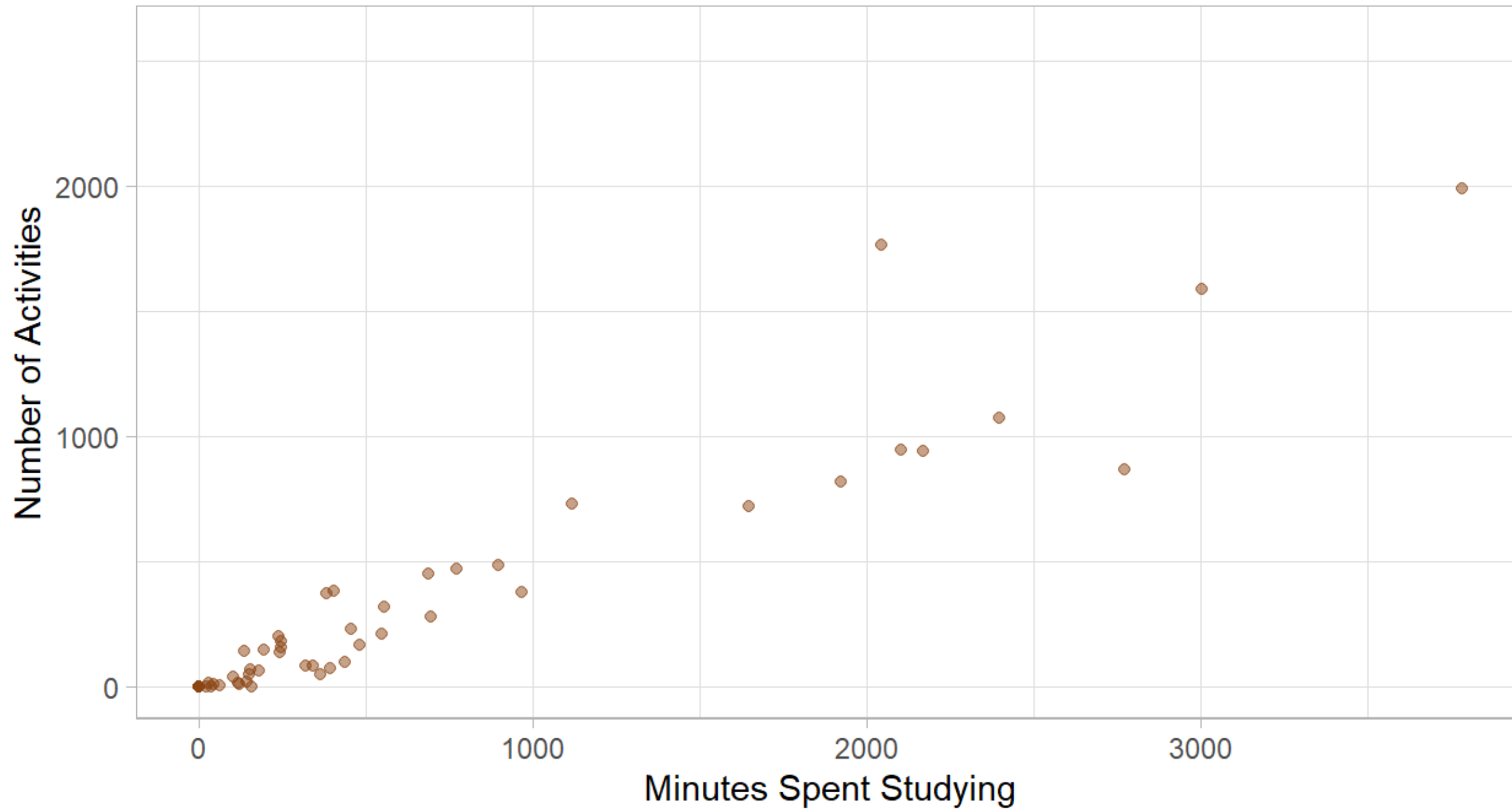
Lessons;

Word Bank;

Grammar Guide.

## Activities Completed vs Study Time

Voxy 2021



## Linear Regression Model:

For every 2-minute increase in study time, 1 activity more is completed (on average).

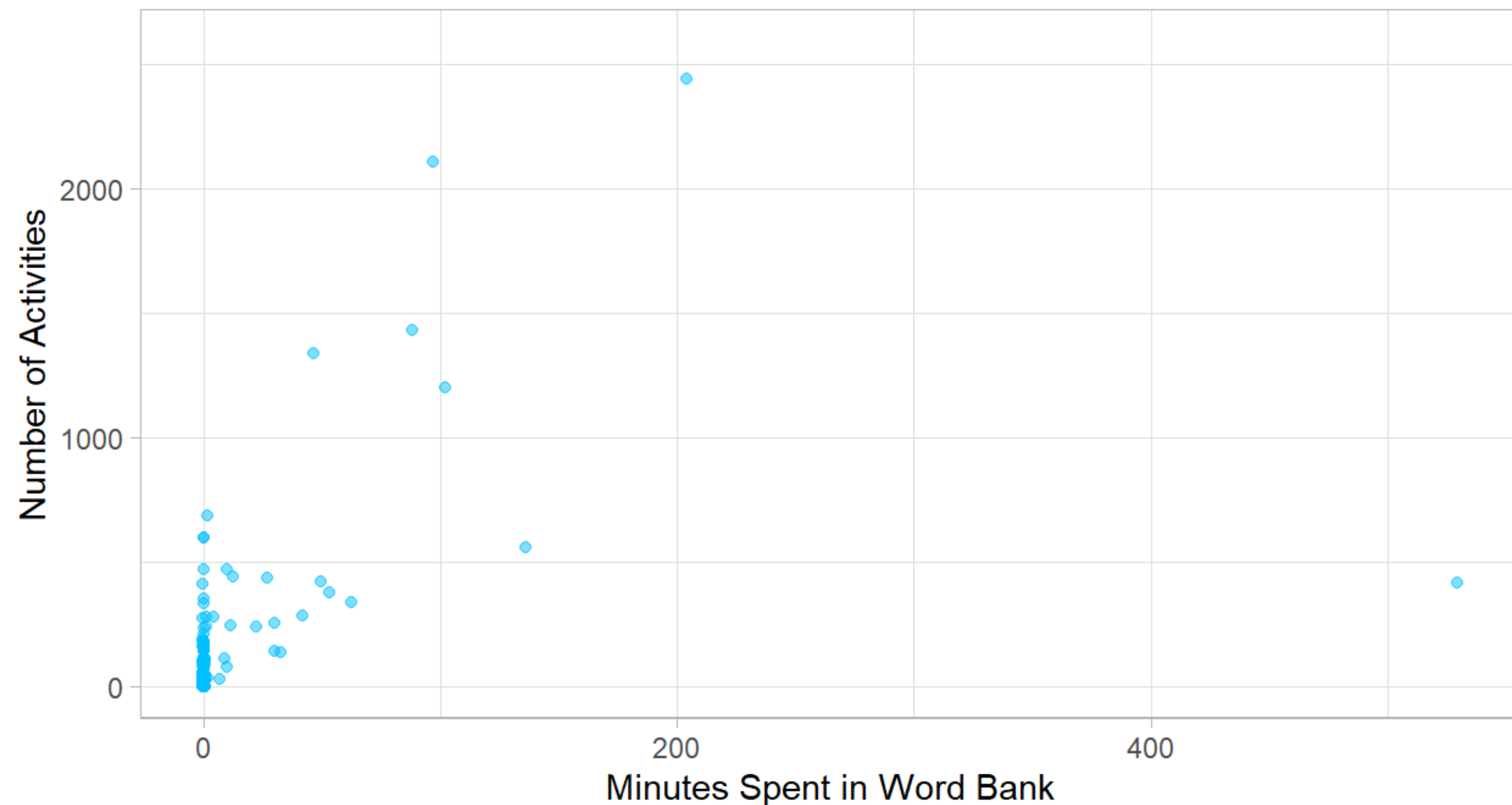
## Activities Completed vs Time in Lessons

Voxy 2021



## Activities Completed vs Time in Word Bank

Voxy 2021



## Linear Regression Model:

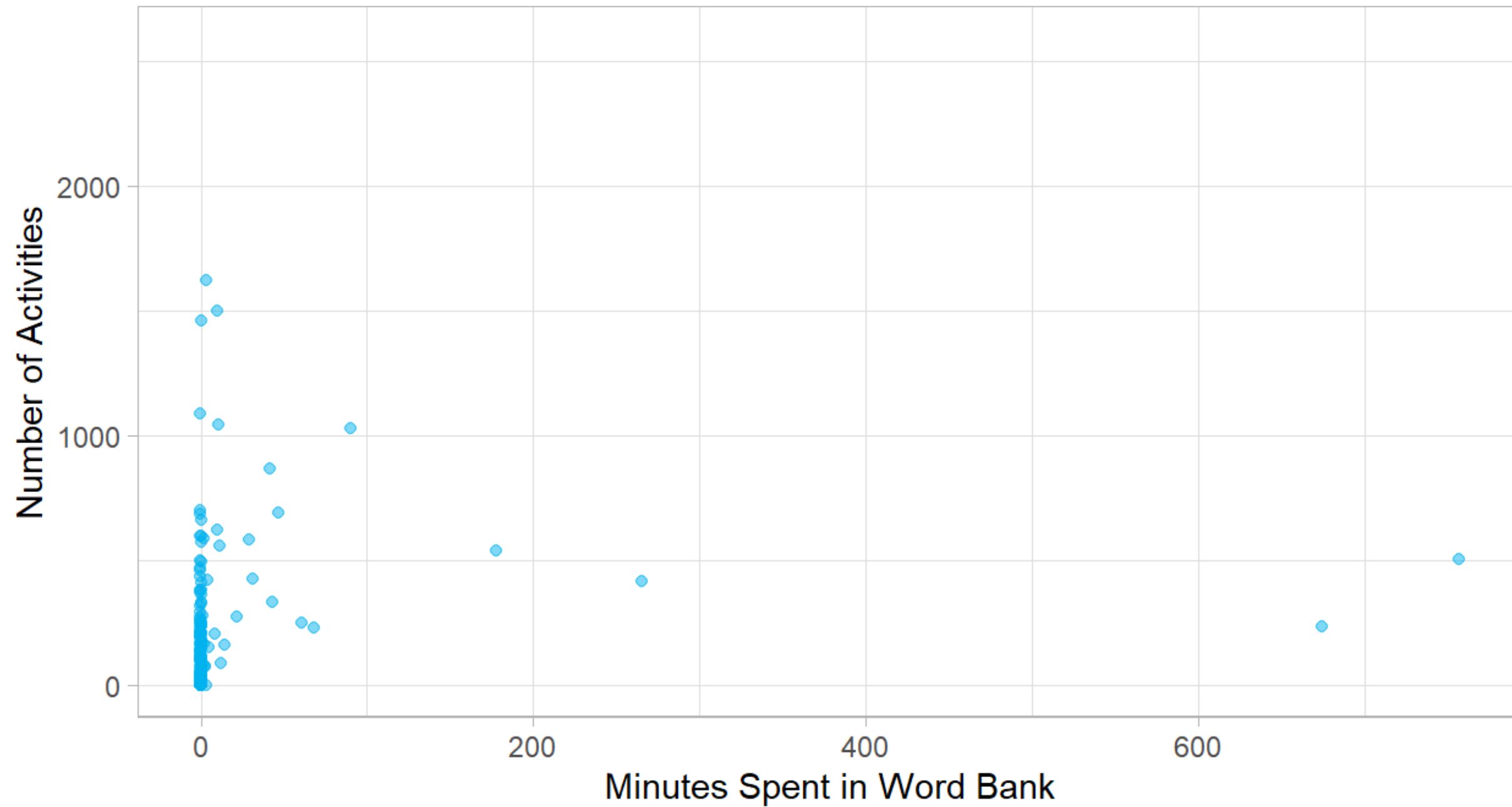
For every 2-minute increase in word bank, almost 7 activities more are completed (on average).

Word Bank really affects activity completion!

Unfortunately many students haven't spent any time in Word Bank.

## Activities Completed vs Time in Grammar Guide

Voxy 2021

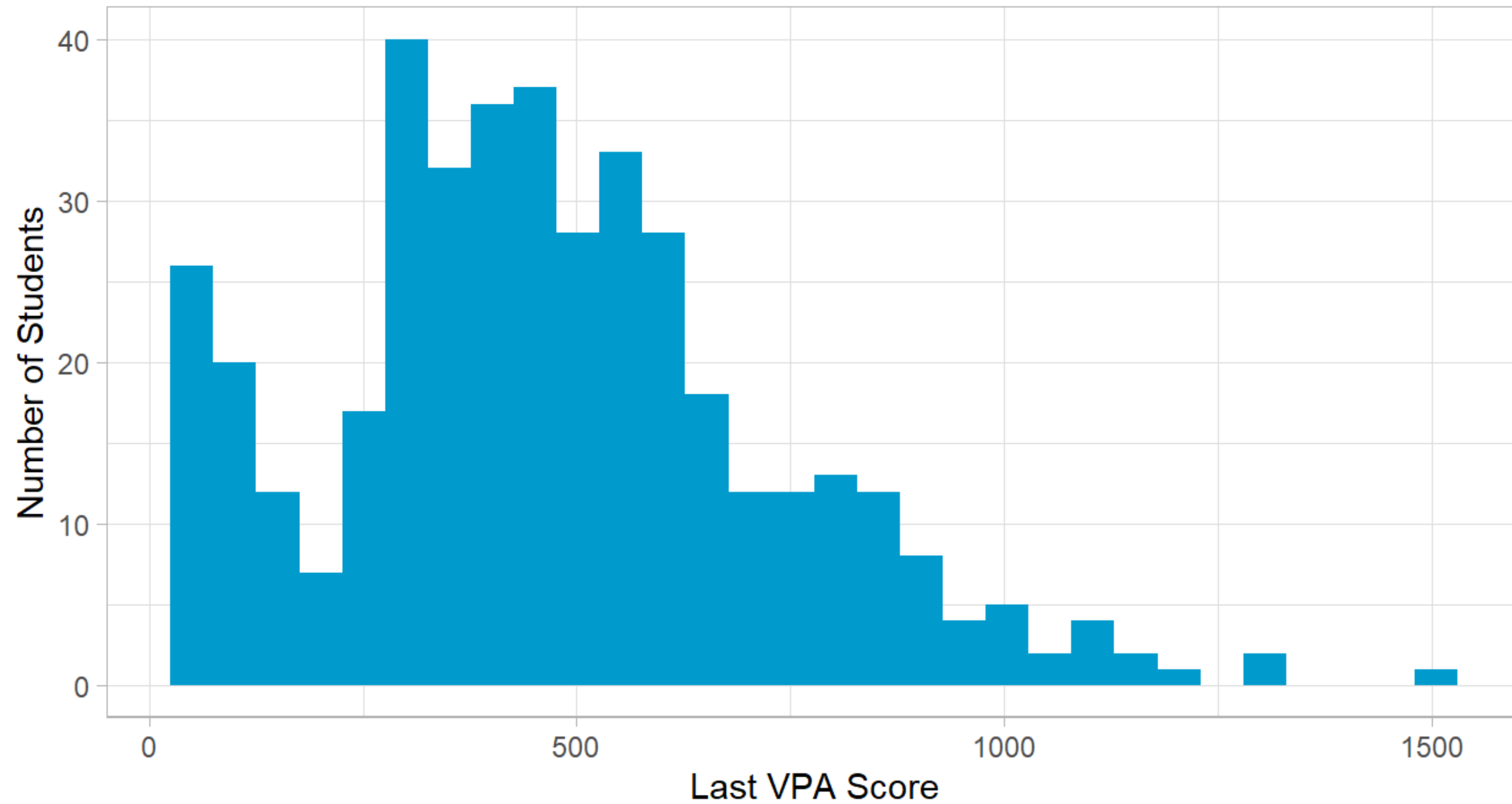


Grammar Guide doesn't seem to affect activity completion.

Most students don't use this tool.

## Distribution of Scores in the Voxy Proficiency Assessment (VPA)

Voxy 2021



Most of the scores are between 300 and 700.

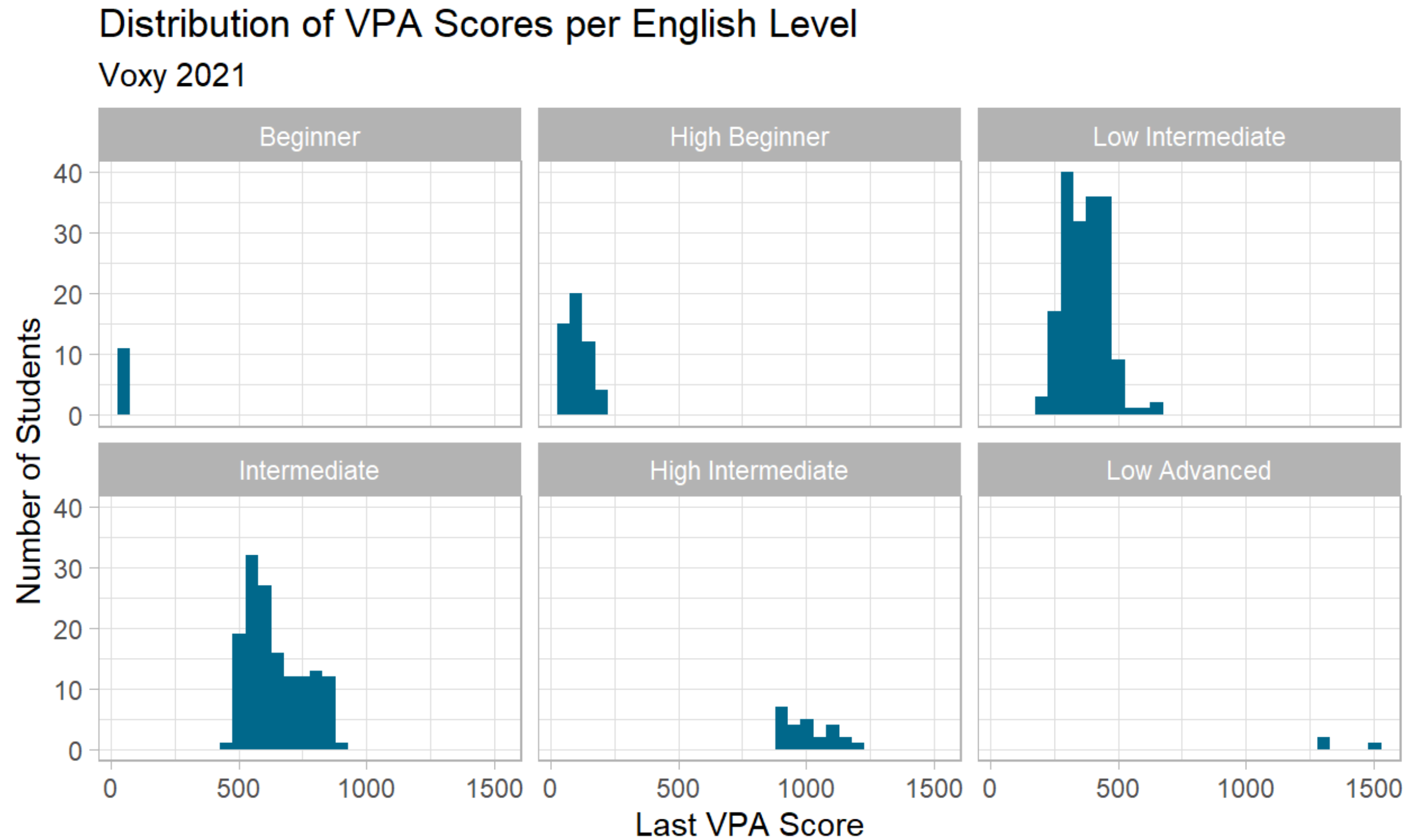
## Summary Statistics for VPA Scores

Mean	Min	1 <sup>st</sup> Quartile	Median	3 <sup>rd</sup> Quartile	Max
473	30	305	445	610	1485

There is a great variability to the vpa results: scores between 30 and 1485!

That's why it is so important to have different English levels that can provide specific aid for each student.

A look at the level distribution based on vpa scores.



This is how Voxy determines the students' English level.



# Thank you!

Juliana Magalhães Rosa