SC1015
Mini Project
HDB Resale Price

FCSA Team 7

Janhavee U2323392K
Vannes Wijaya U2320201J
Thwun Thiri Thu U2321893B

01

Introduction

Practical Motivations

- HDB flats are home to the majority of people in Singapore (77.8%)*
- Implications for buyers, sellers as well as policy makers in Singapore

*)source: https://smartwealth.sg/housing-household-statistics-singapore/

Problem Formulation

1

Which variables has the greatest influence in HDB resale price?

2

Which model would be ideal for HDB resale price prediction?

3

How are HDB resales distributed across Singapore?

02

Data Preparation

Datasets

Singapore HDB Flat Resale Prices (1990-2020)

31 | New Notebook | Download (9 MB)

Data Card | Code (5) | Discussion (0) | Suggestions (0)

month year and month of transaction	town town of flat	flat_type type of flat	block block of flat	street_name street name of flat	storey_r storey rang
2017-01-01 2020-09-01	SENGKANG 8% WOODLANDS 7% Other (67947) 85%	4 ROOM 41% 5 ROOM 25% Other (27145) 34%	2446 unique values	YISHUN RING RD 2% BEDOK RESERVOIR ... 1% Other (78194) 97%	04 TO 06 07 TO 09 Other (447
2017-01	ANG MO KIO	3 ROOM	534	ANG MO KIO AVE 10	01 TO 03
2017-01	ANG MO KIO	3 ROOM	233	ANG MO KIO AVE 3	10 TO 12
2017-01	ANG MO KIO	3 ROOM	235	ANG MO KIO AVE 3	04 TO 06
2017-01	ANG MO KIO	3 ROOM	219	ANG MO KIO AVE 1	07 TO 09
2017-01	ANG MO KIO	3 ROOM	536	ANG MO KIO AVE 10	07 TO 09
2017-01	ANG MO KIO	3 ROOM	230	ANG MO KIO AVE 3	04 TO 06
2017-01	ANG MO KIO	3 ROOM	570	ANG MO KIO AVE 3	10 TO 12
2017-01	ANG MO KIO	3 ROOM	624	ANG MO KIO AVE 4	04 TO 06
2017-01	ANG MO KIO	3 ROOM	441	ANG MO KIO AVE 10	07 TO 09
2017-01	ANG MO KIO	3 ROOM	625	ANG MO KIO AVE 9	04 TO 06
2017-01	ANG MO KIO	3 ROOM	119	ANG MO KIO AVE 3	07 TO 09
2017-01	ANG MO KIO	3 ROOM	255	ANG MO KIO AVE 4	04 TO 06
2017-01	ANG MO KIO	3 ROOM	432	ANG MO KIO AVE 10	10 TO 12

resale-flat-prices-based-on-r
resale-flat-prices-based-on-r
resale-flat-prices-based-on-r

Summary

5 files

52 columns

https://www.kaggle.com/datasets/teyang/singapore-hdb-flat-resale-prices-19902020

Data preparation

Merge datasets from 1990 to 2020

Remove invalid and duplicated entries

Adjust variables format:

- Extract RESALE_YEAR from the transaction date
- Encode STOREY_RANGE into numeric

Split dataset into train and test set

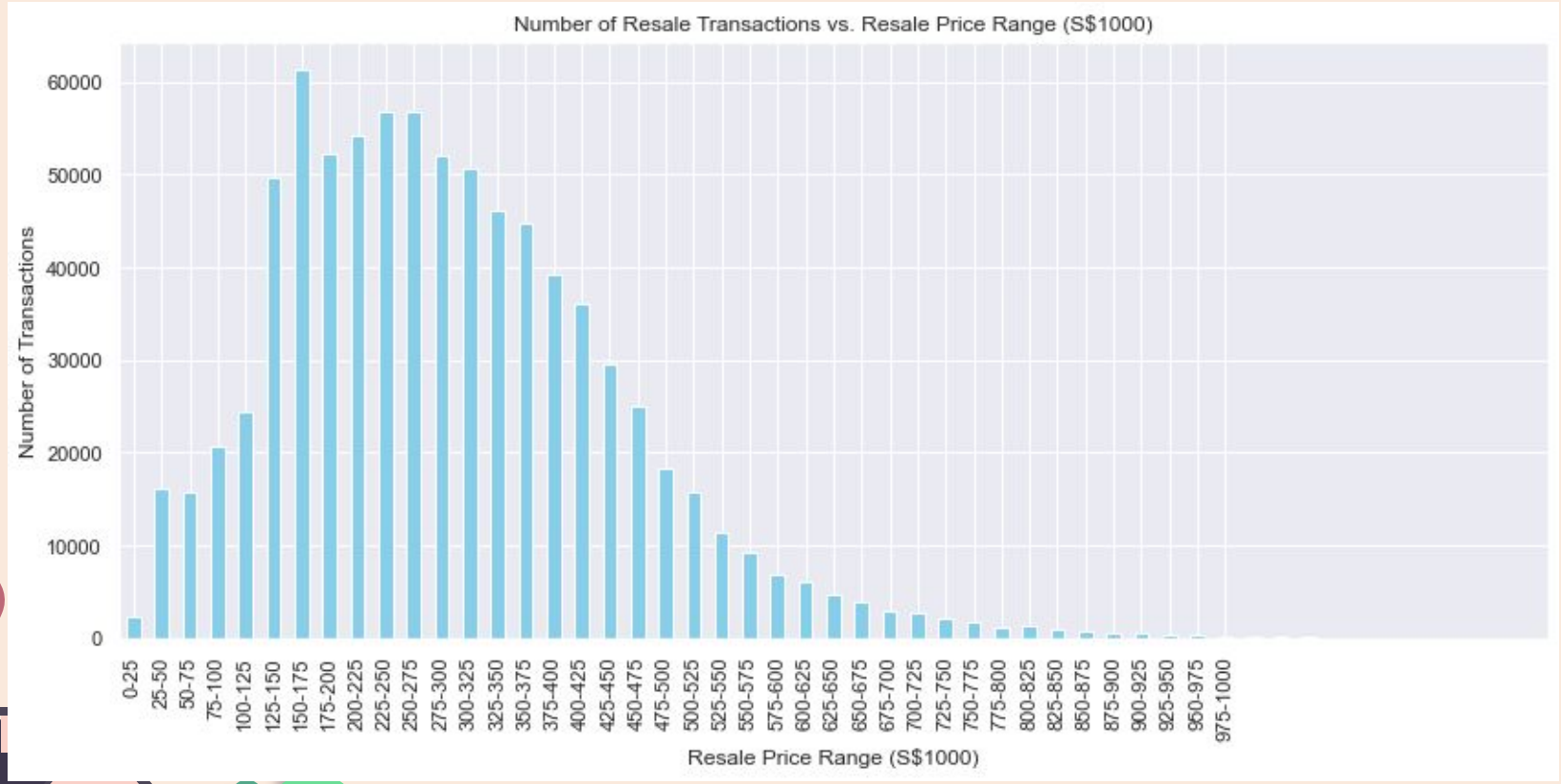
Remove outliers from train and test set

03

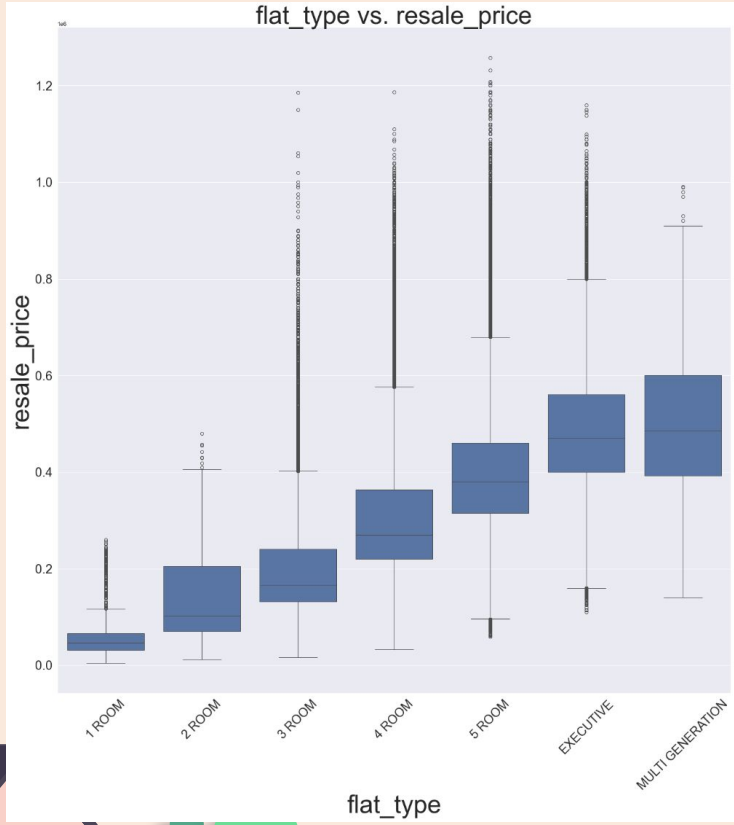
Exploratory Data Analysis

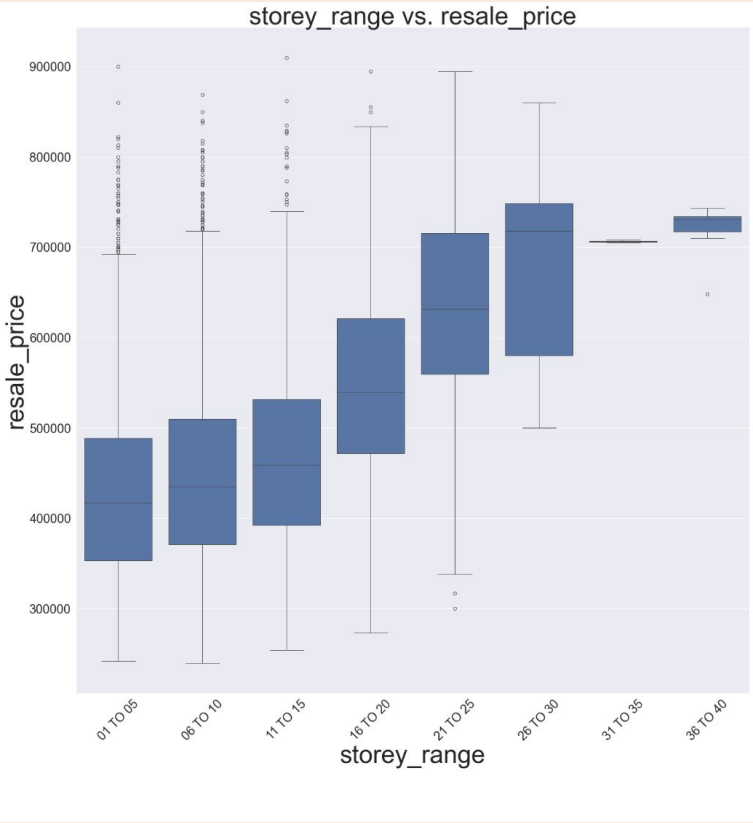
- Bar plot
- Box plot
- Scatter plot
- Time series
- Heatmap

Data Visualization

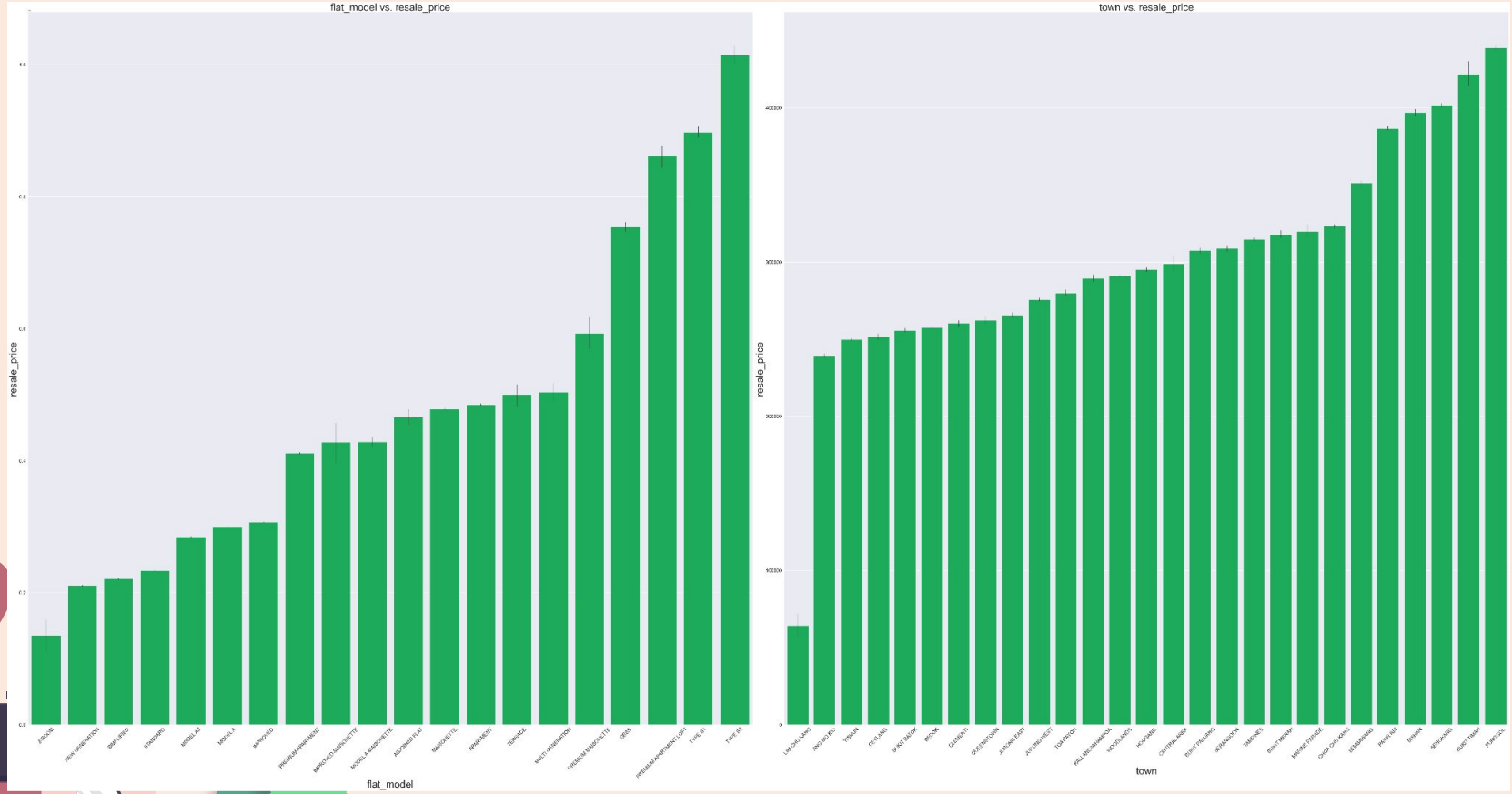


Data Visualization: Categorical

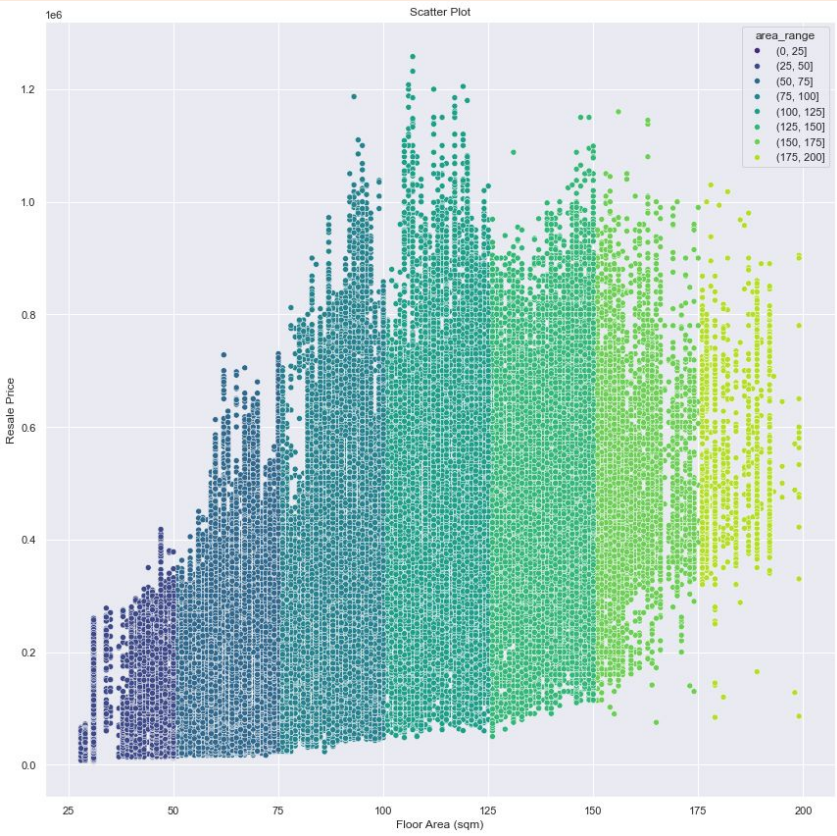


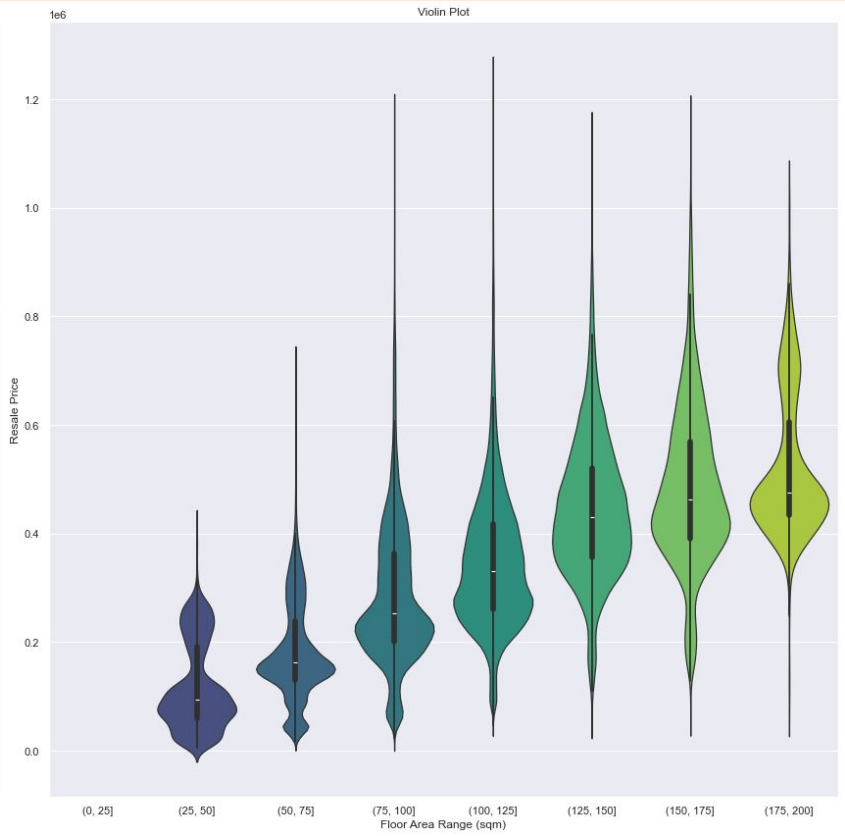


Data Visualization: Categorical

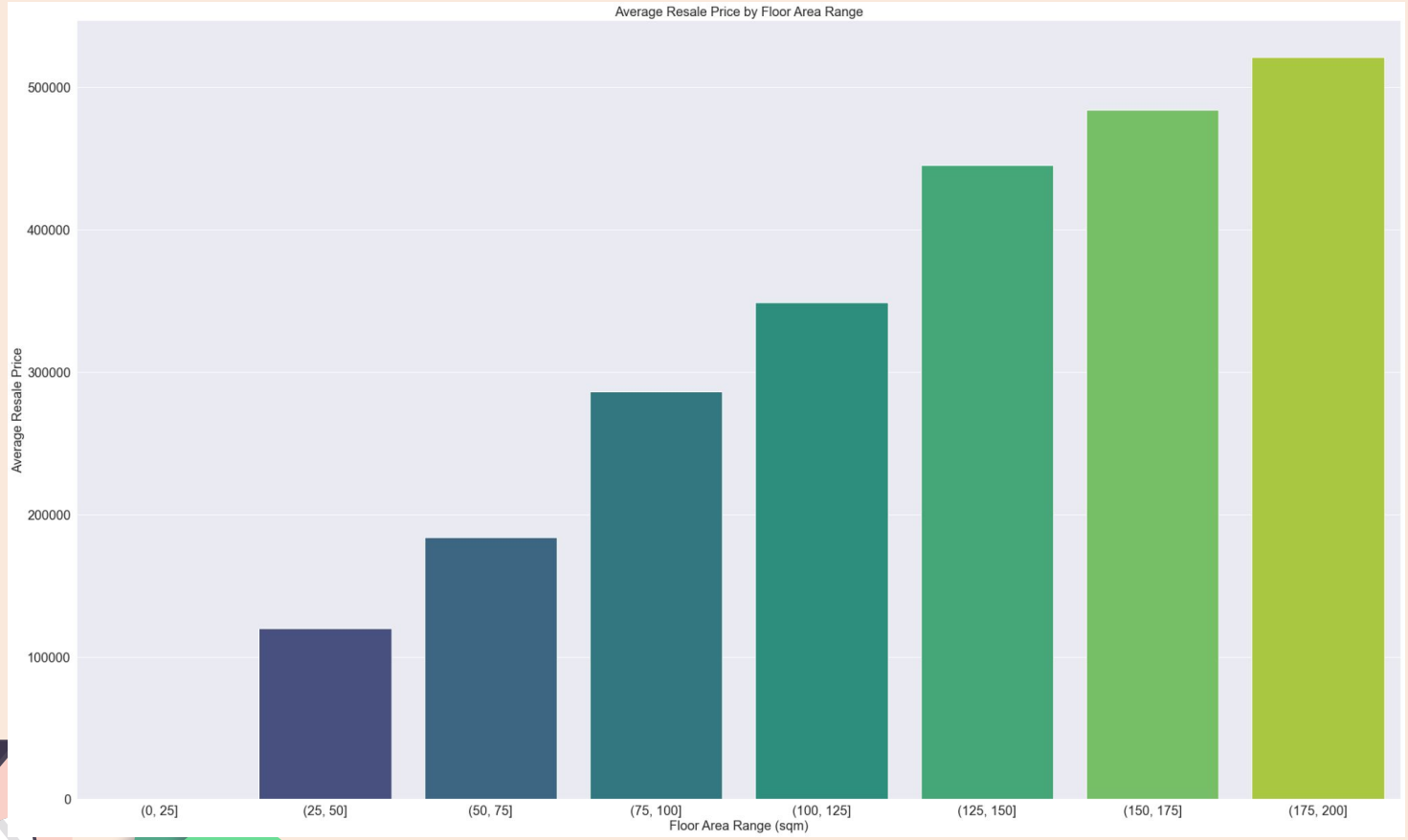


Data Visualization: Numerical

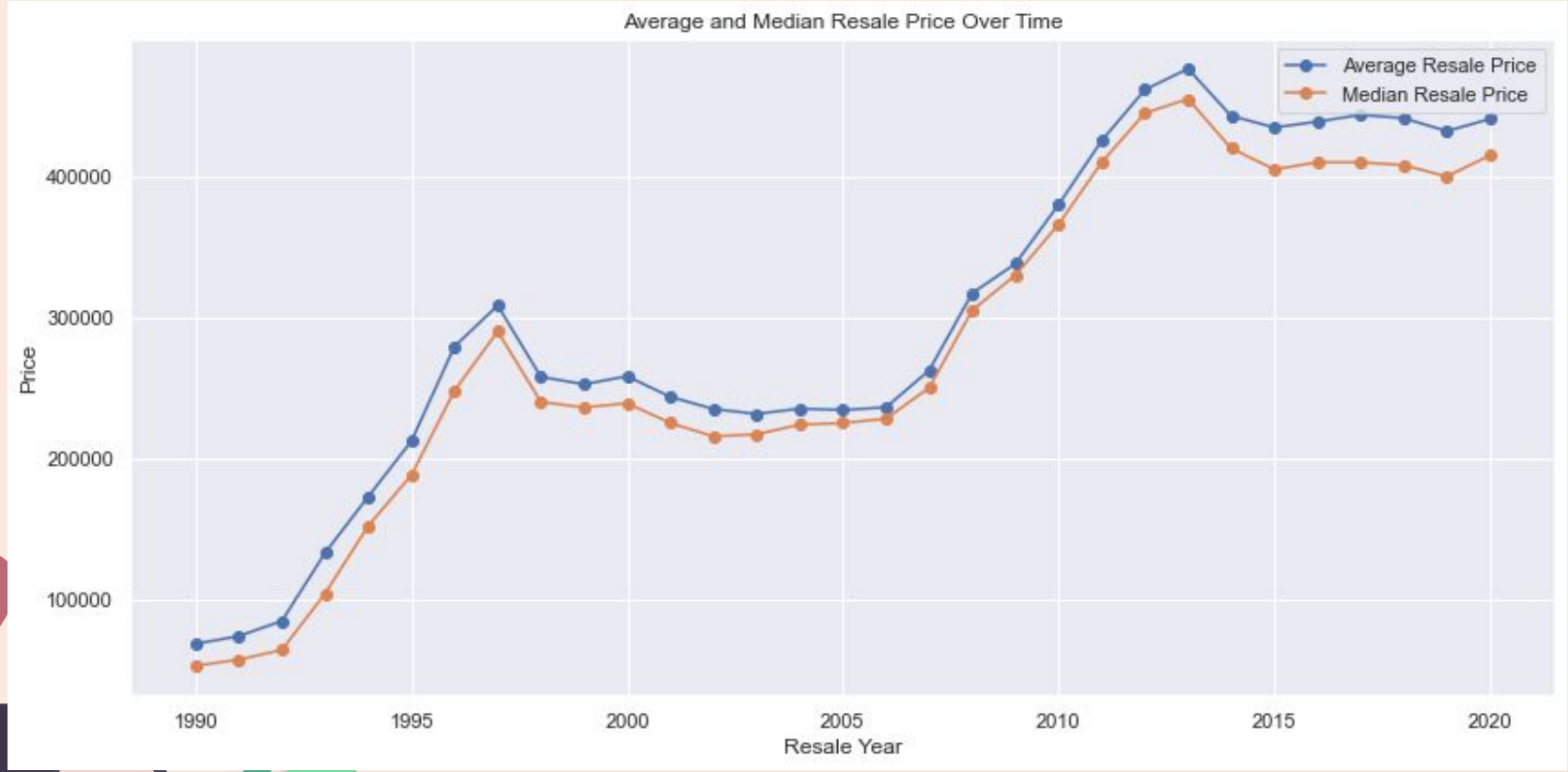




Data Visualization: Numerical



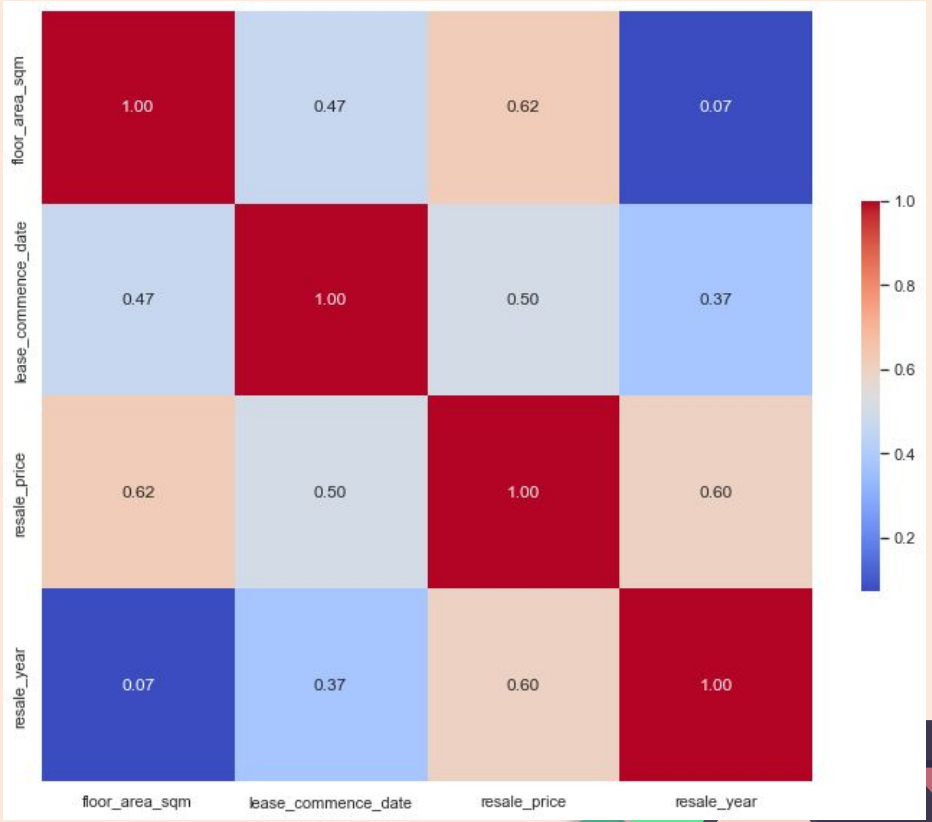
Data Visualization: Time series



Data Visualization: Heatmap

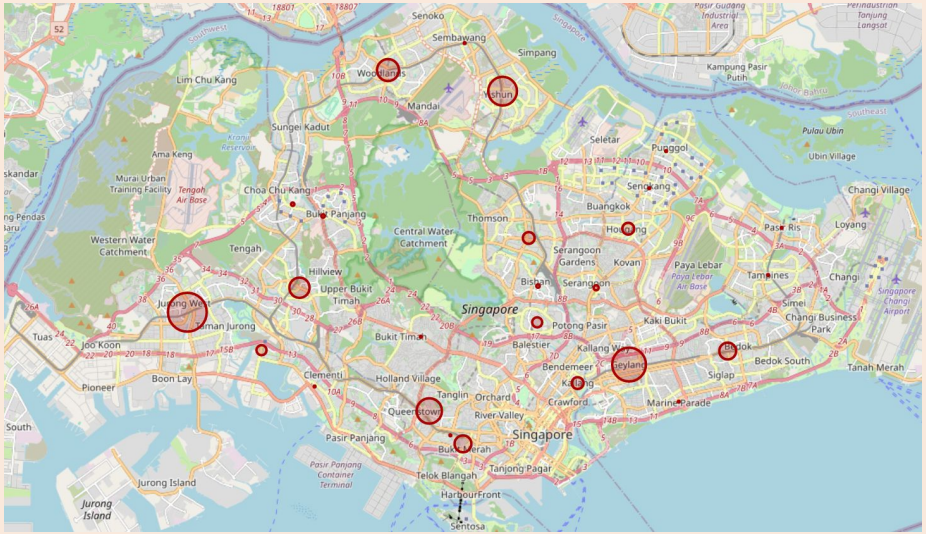
In order of correlation with resale_price:

1. **floor_area_sqm**
2. **resale_year**
3. **lease_commence_date**

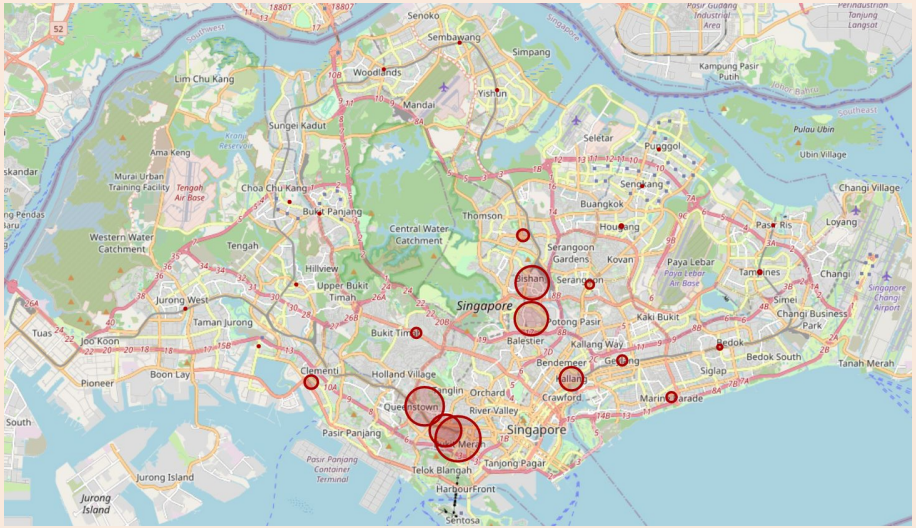


Data Visualization: Interactive map

Lowest 1% Resale_price, 2005-2020



Highest 1% Resale_price, 2005-2020



04

Machine Learning

Predictors

1) Floor area

2) Storey range

3) Lease commence date

4) Resale date

Machine Learning Models

1) Linear Regression

2) Ridge Regression

3) Bayesian Ridge Regression

4) Gradient Boosting

5) Random Forest

Machine Learning

R^2 Predictor \ Model	Linear Regression	Ridge Regression	Bayesian Ridge	Gradient Boosting	Random Forest
Floor Area	0.411	0.411	0.411	0.449	0.451
Lease Date	0.246	0.246	0.246	0.275	0.275
Storey Range	0.022	0.022	0.022	0.046	0.046
Resale Year	0.356	0.356	0.356	0.470	0.470
All	0.736	0.736	0.736	0.884	0.905

05

Data-driven Insights

Which variables has the greatest influence in resale price?

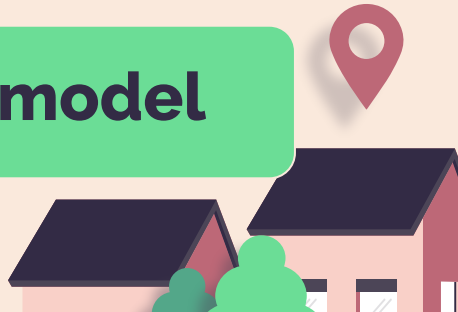
What we look for:

- Distinction in visualization
- High correlation in Heatmap
- Consistent good result in ML models
  - Low Mean Squared Error
  - High Explained Variance

Floor area

Resale year

Flat model



Which model would be ideal for HDB resale price prediction?

What we look for:

- Low Mean Squared Error
- Low Root Mean Squared Error
- High Explained Variance

Random Forest

RMSE = S$41,783

Gradient Boosting

RMSE = S$46,240

How are HDB resales distributed across Singapore?

Interactive Map →

Thank You

Linear Regression

```
Linear Regression using ['floor_area_sqm']:
Mean Squared Error (MSE): 10852322747.407
R-squared Score: 0.41058981333644384


Linear Regression using ['lease_commence_date']:
Mean Squared Error (MSE): 13890149818.691706
R-squared Score: 0.2455996759425535


Linear Regression using ['storey_range']:
Mean Squared Error (MSE): 17999438190.340294
R-squared Score: 0.02241644756257466


Linear Regression using ['resale_year']:
Mean Squared Error (MSE): 11860717498.754131
R-squared Score: 0.355821985982242


Linear Regression using ['floor_area_sqm', 'lease_commence_date', 'storey_range', 'resale_year']:
Mean Squared Error (MSE): 4869618849.629878
R-squared Score: 0.7355217844192378
```

Ridge Regression

```
Ridge Regression using ['floor_area_sqm']:
Mean Squared Error (MSE): 10852322747.584032
R-squared Score: 0.41058981332682887


Ridge Regression using ['lease_commence_date']:
Mean Squared Error (MSE): 13890149818.726665
R-squared Score: 0.24559967594065468


Ridge Regression using ['storey_range']:
Mean Squared Error (MSE): 17999438194.853462
R-squared Score: 0.022416447317455956


Ridge Regression using ['resale_year']:
Mean Squared Error (MSE): 11860717502.674099
R-squared Score: 0.3558219857693412


Ridge Regression using ['floor_area_sqm', 'lease_commence_date', 'storey_range', 'resale_year']:
Mean Squared Error (MSE): 4869618852.099425
R-squared Score: 0.7355217842851121
```

Bayesian Ridge Regression

```
Bayesian Ridge Regression using ['floor_area_sqm']:
Mean Squared Error (MSE): 10852322917.217775
R-squared Score: 0.4105898041136993


Bayesian Ridge Regression using ['lease_commence_date']:
Mean Squared Error (MSE): 13890149828.021973
R-squared Score: 0.24559967543580896


Bayesian Ridge Regression using ['storey_range']:
Mean Squared Error (MSE): 17999439435.067406
R-squared Score: 0.022416379959089627


Bayesian Ridge Regression using ['resale_year']:
Mean Squared Error (MSE): 11860717947.097328
R-squared Score: 0.3558219616318745


Bayesian Ridge Regression using ['floor_area_sqm', 'lease_commence_date', 'storey_range', 'resale_year']:
Mean Squared Error (MSE): 4869619217.818021
R-squared Score: 0.7355217644222429
```

Gradient Boosting

```
Gradient Boosting Regression using ['floor_area_sqm']:
Mean Squared Error: 10147424080.993206
R-squared Score: 0.44887419394511274


Gradient Boosting Regression using ['lease_commence_date']:
Mean Squared Error: 13351409455.82496
R-squared Score: 0.2748596846274547


Gradient Boosting Regression using ['storey_range']:
Mean Squared Error: 17571429926.24127
R-squared Score: 0.04566238640054632


Gradient Boosting Regression using ['resale_year']:
Mean Squared Error: 9755997498.846968
R-squared Score: 0.47013331240461653


Gradient Boosting Regression using ['floor_area_sqm', 'lease_commence_date', 'storey_range', 'resale_year']:
Mean Squared Error: 2138139236.7484896
R-squared Score: 0.8838736115781558
```

Random Forest

```
Random Forest Regression using ['floor_area_sqm']:
Mean Squared Error: 10110927498.159582
R-squared Score: 0.45085639242936715


Random Forest Regression using ['lease_commence_date']:
Mean Squared Error: 13346982729.891596
R-squared Score: 0.2751001084905689


Random Forest Regression using ['storey_range']:
Mean Squared Error: 17571357488.52167
R-squared Score: 0.04566632063018761


Random Forest Regression using ['resale_year']:
Mean Squared Error: 9755938660.965277
R-squared Score: 0.4701365080013231


Random Forest Regression using ['floor_area_sqm', 'lease_commence_date', 'storey_range', 'resale_year']:
Mean Squared Error: 1745874211.190154
R-squared Score: 0.9051782674861425
```

CASE STUDIES

INSTAGRAM GIVEAWAY

Giveaway of a product for our second anniversary to increase the followers

APPROACH

- Preparing a set of posts and stories we can share on Instagram to let more people know about our brand
- We can get shared more and mabe reach virality if our content is attractive and easy to share

RESULTS

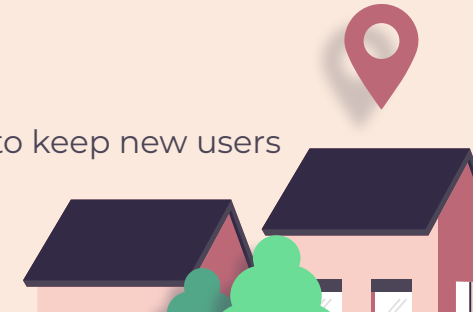
23%	80%	+10%
Increase in shares	Increase in followers	Increase in clicks

KEY TAKEAWAY 1

Content posted on weekdays is more shared

KEY TAKEAWAY 2

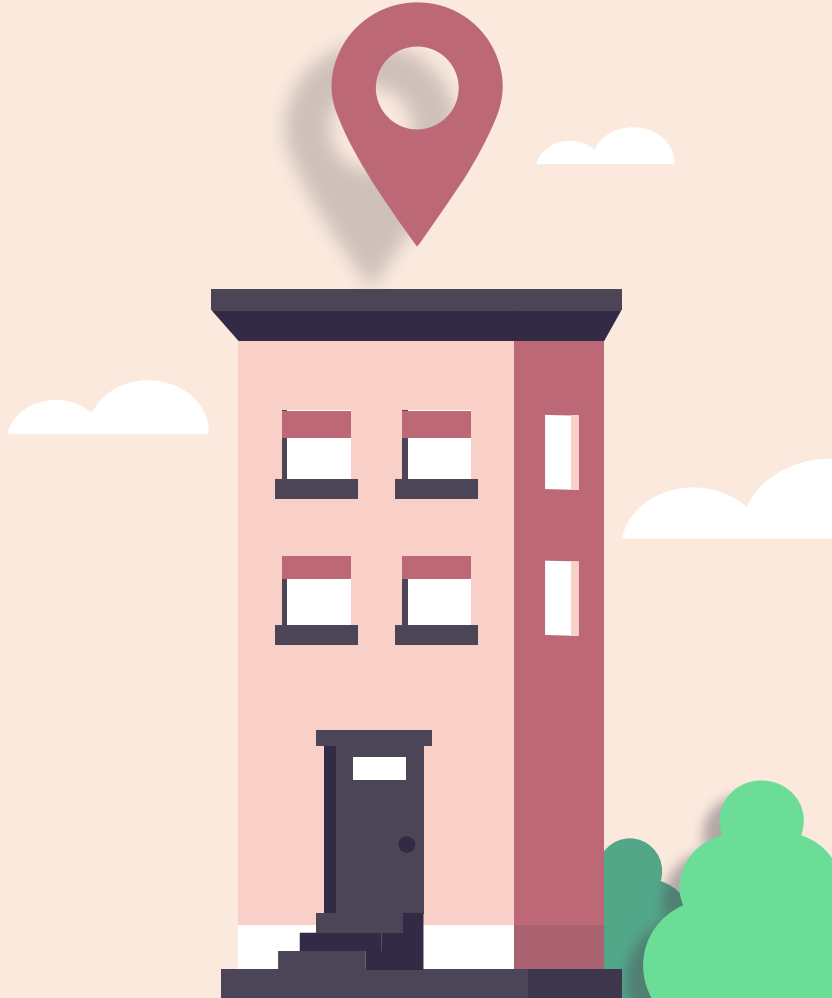
We must design new content to keep new users



THANKS

DOES ANYONE HAVE ANY QUESTIONS?

youremail@freepik.com
+91 620 421 838
yourwebsite.com

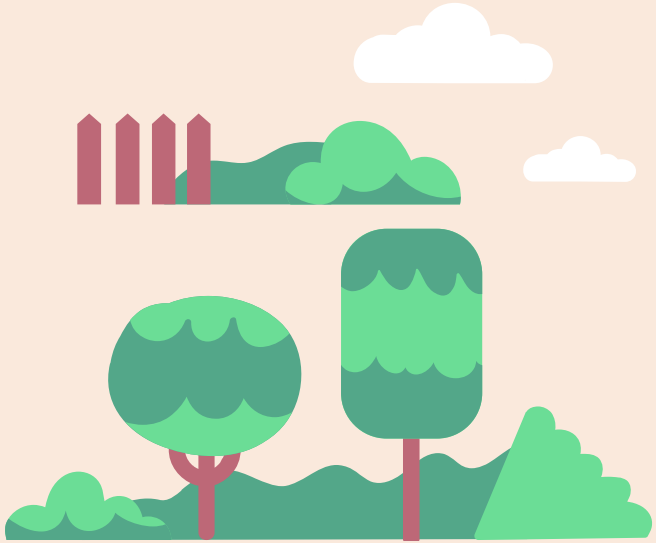


ICON PACK

ALTERNATIVE RESOURCES

Here’s an assortment of alternative resources whose style fits the one of this template:

- Collection of creative different houses
- House for sale / for rent



RESOURCES

Did you like the resources on this template? Get them for free at our other websites:

VECTORS

- Gradient real estate facebook cover
- Gradient real estate landing page
- Gradient real estate linkedin banner

ICONS

- Icon Pack: Real Estate

PHOTOS

- Realtor handing smiley couple the keys for their new home
- Front view of beautiful woman with copy space

Instructions for use

If you have a free account, in order to use this template, you must credit **Slidesgo** by keeping the **Thanks** slide. Please refer to the next slide to read the instructions for premium users.

As a Free user, you are allowed to:

- Modify this template.
- Use it for both personal and commercial projects.

You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs or visit our blog:
https://slidesgo.com/faqs and **https://slidesgo.com/slidesgo-school**

Instructions for use (premium users)

As a Premium user, you can use this template without attributing **Slidesgo** or keeping the **Thanks** slide.

You are allowed to:

- Modify this template.
- Use it for both personal and commercial purposes.
- Hide or delete the “Thanks” slide and the mention to Slidesgo in the credits.
- Share this template in an editable format with people who are not part of your team.

You are not allowed to:

- Sublicense, sell or rent this Slidesgo Template (or a modified version of this Slidesgo Template).
- Distribute this Slidesgo Template (or a modified version of this Slidesgo Template) or include it in a database or in any other product or service that offers downloadable images, icons or presentations that may be subject to distribution or resale.
- Use any of the elements that are part of this Slidesgo Template in an isolated and separated way from this Template.
- Register any of the elements that are part of this template as a trademark or logo, or register it as a work in an intellectual property registry or similar.

For more information about editing slides, please read our FAQs or visit our blog:
https://slidesgo.com/faqs and **https://slidesgo.com/slidesgo-school**

Fonts & colors used

This presentation has been made using the following fonts:

Raleway
(https://fonts.google.com/specimen/Raleway)

Montserrat
(https://fonts.google.com/specimen/Montserrat)

Storyset

Create your Story with our illustrated concepts. Choose the style you like the most, edit its colors, pick the background and layers you want to show and bring them to life with the animator panel! It will boost your presentation. Check out **how it works**.

Pana

Amico

Bro

Rafiki

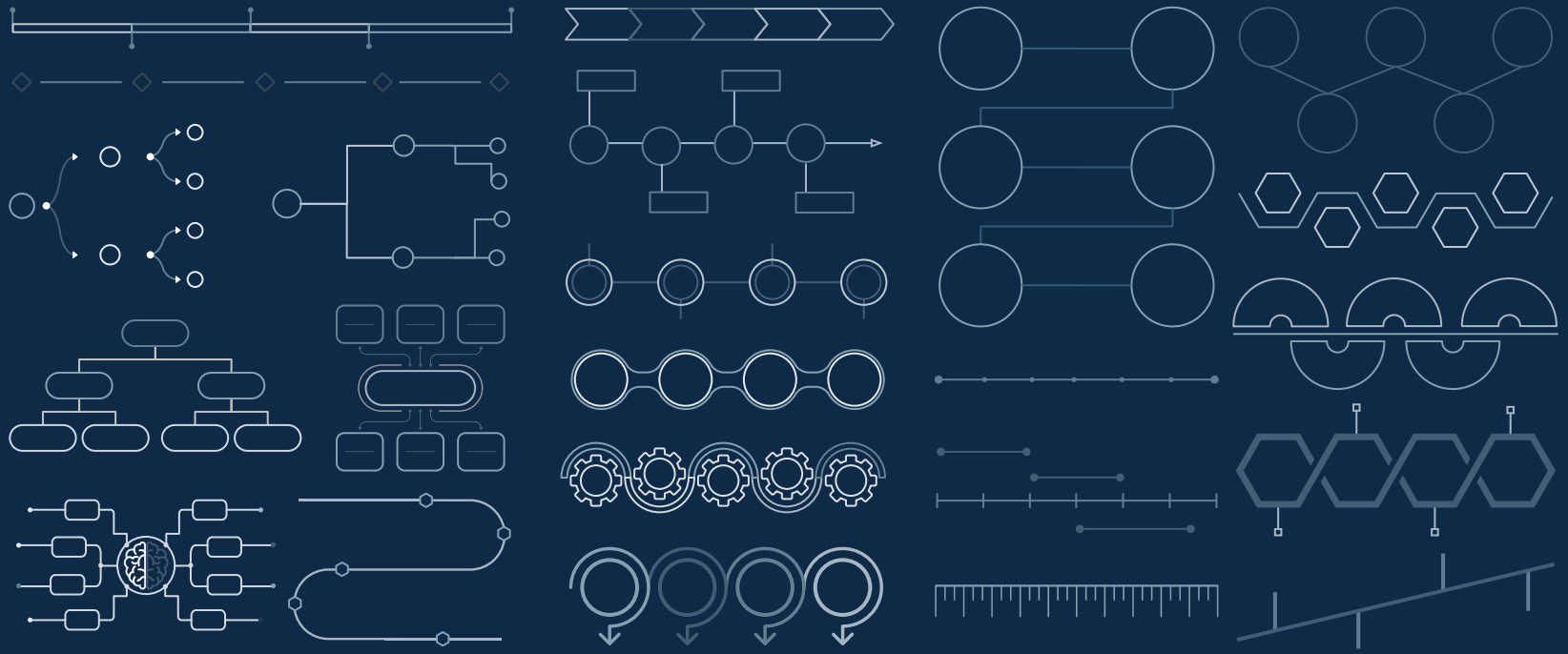
Cuate

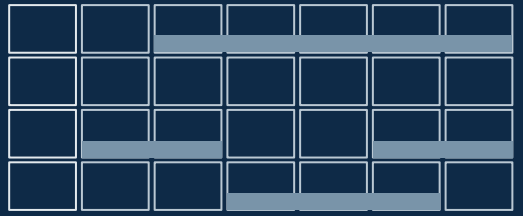
Use our editable graphic resources...

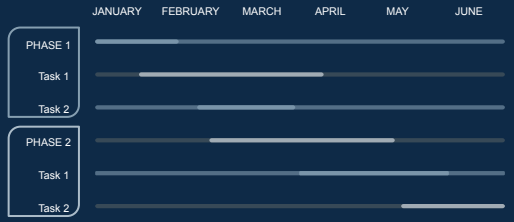
You can easily **resize** these resources without losing quality. To **change the color**, just ungroup the resource and click on the object you want to change. Then, click on the paint bucket and select the color you want. Group the resource again when you're done. You can also look for more **infographics** on Slidesgo.

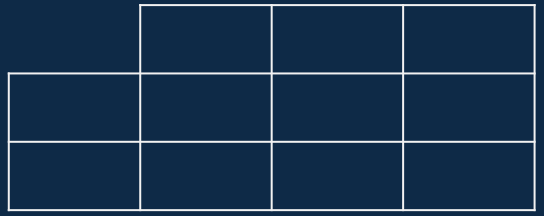


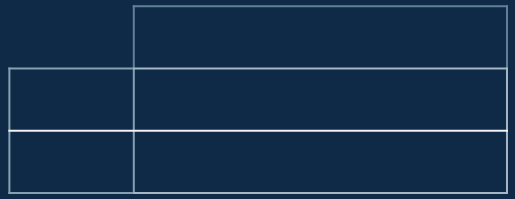


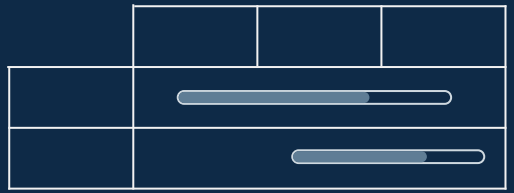



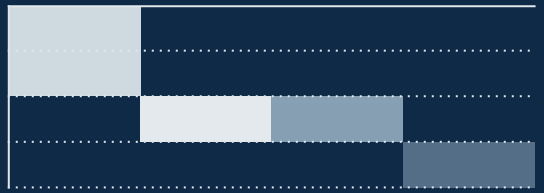


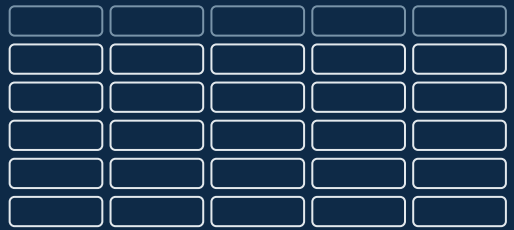

JANUARY
FEBRUARY
MARCH
APRIL
MAY
JUNE
PHASE 1
Task 1
Task 2
PHASE 2
Task 1
Task 2

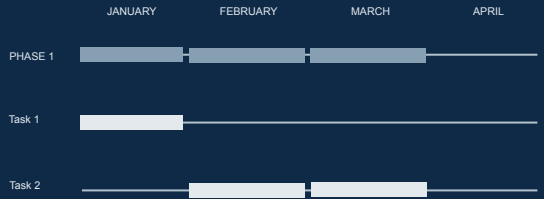


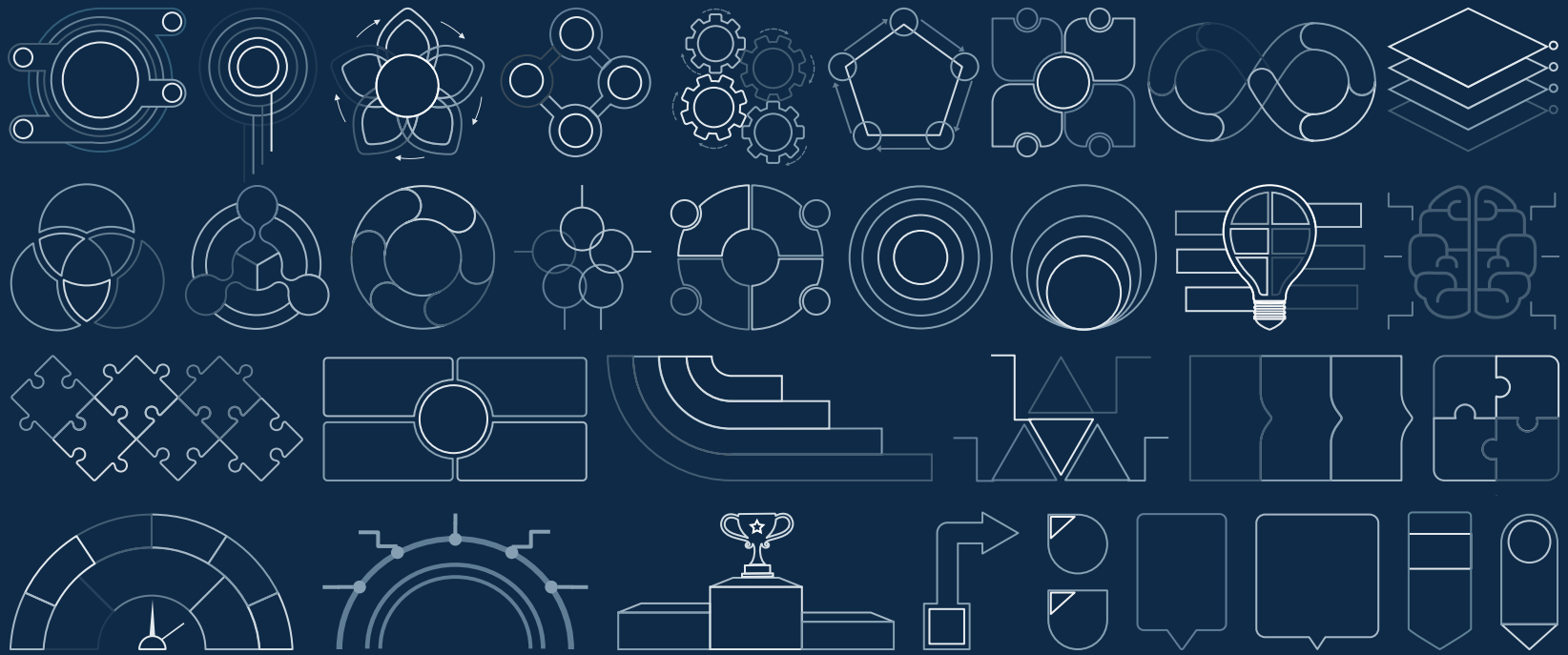


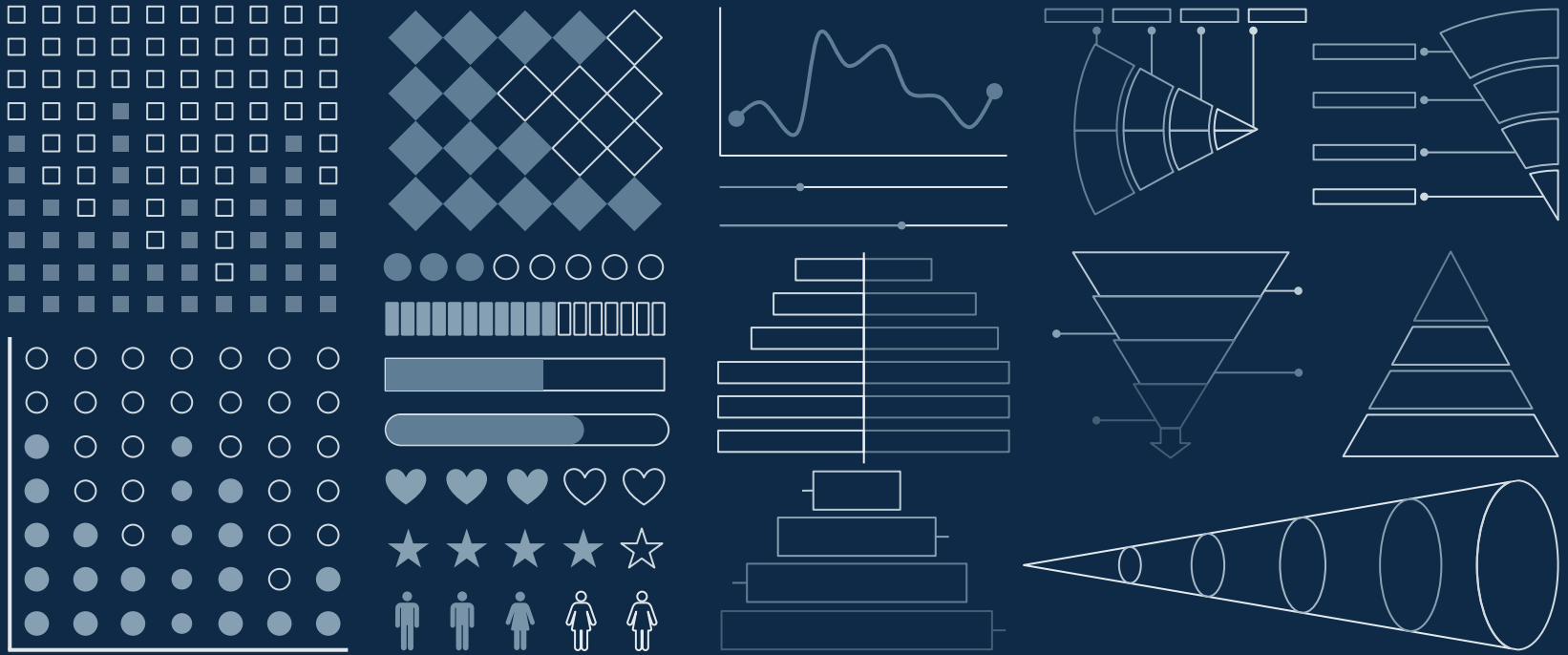







JANUARY
FEBRUARY
MARCH
APRIL
PHASE 1
Task 1
Task 2





...and our sets of editable icons

You can **resize** these icons without losing quality.
You can **change the stroke and fill color**; just select the icon and click on the **paint bucket/pen**.
In Google Slides, you can also use **Flaticon's extension**, allowing you to customize and add even more icons.

Educational Icons

Medical Icons

Business Icons

Teamwork Icons

Help & Support Icons

Avatar Icons

Creative Process Icons

Performing Arts Icons

Nature Icons

SEO & Marketing Icons


slidesgo