Tahoe Healthcare Systems

Can patient education save money?

Jonathan Hilgart

Agenda

- Overview of Tahoe Healthcare
- Can we use a ‘quick’ heuristic (managerial clusters) to find people who will be readmitted?
- Does a machine learning model add more value to predict readmission?
- How much money could Tahoe Healthcare save?
- Recommendations

Overview of Tahoe Healthcare

Problem Statement & Variables

- Readmissions penalties cost $750,000 in 2012
- <u>Assumptions:</u>
  - $8k penalty per readmit
  - $1.2k cost per caretracker
  - Caretracker has 40% effectiveness

- Age of patient
- In flu season or not
- Severity score
- Comorbidity score
- Gender
- Emergency room admit
- Indicator for readmission after 30 days
- ~4,300 patients

Managerial Clusters

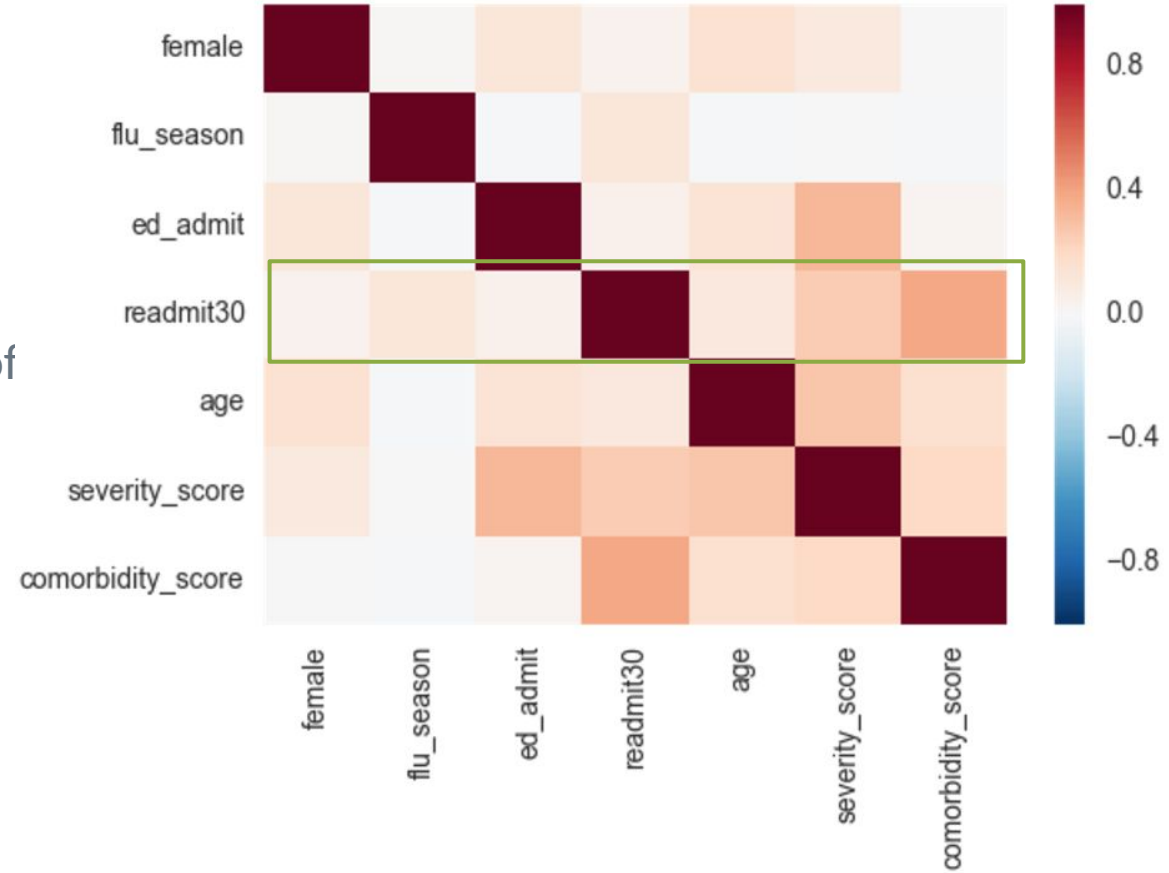
A ‘quick’ heuristic

Correlation with readmission (readmit30)

1) Comorbidty score (.38)
2) Severity score (.25)
3) Flu season (.12)

- Create 32 managerial clusters from the quantiles of these factors

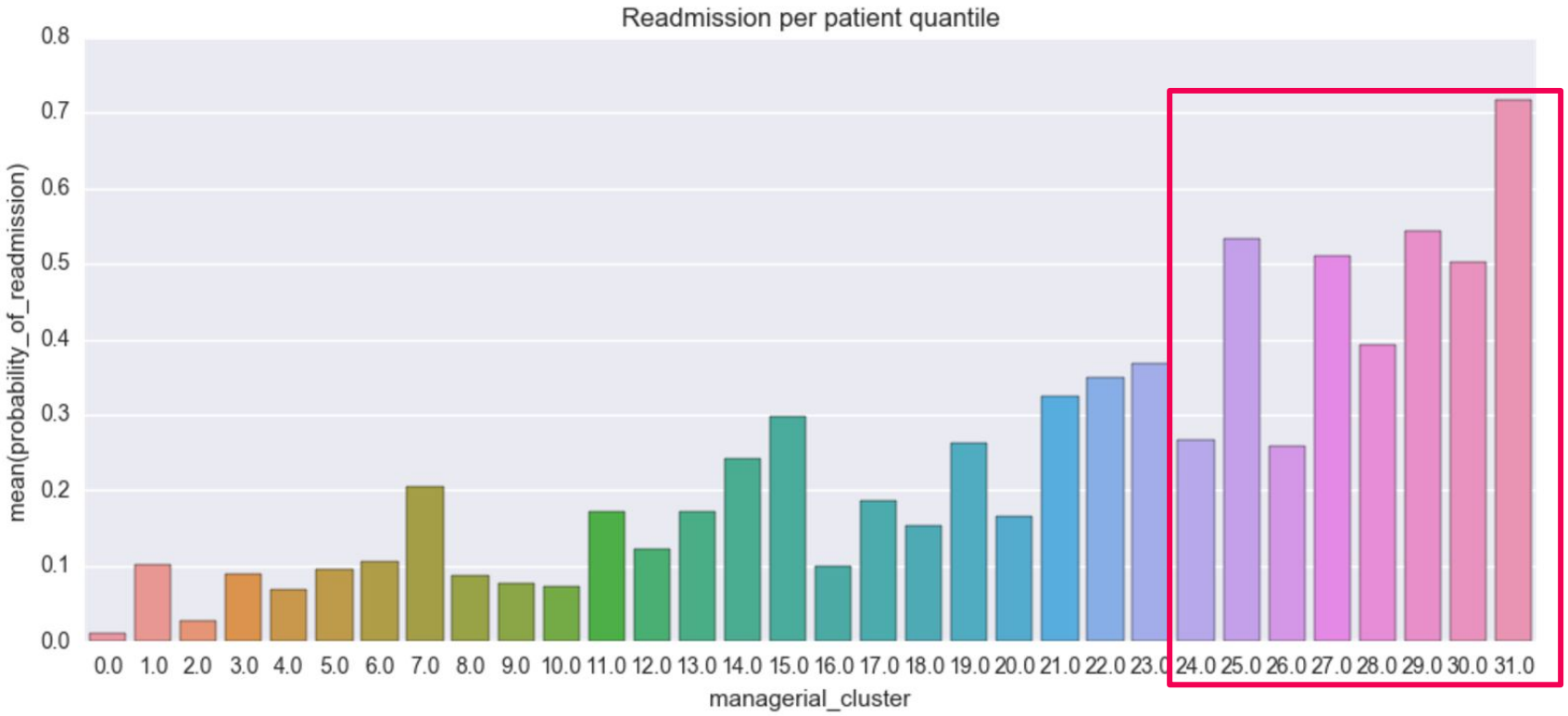
	severity_score	comorbidity_score	flu_season
count	4382.000000	4382.000000	4382.000000
mean	22.353263	94.735509	0.408033
std	18.005320	57.145155	0.491525
min	1.000000	1.000000	0.000000
25%	8.000000	51.000000	0.000000
50%	19.000000	86.000000	0.000000
75%	32.000000	131.000000	1.000000
max	112.000000	322.000000	1.000000



**severity score : health score based on lab tests and vital signs*
**comorbodity score: health score based upon pre-existing patient conditions*

Managerial Cluster Analysis

- Baseline readmission rate ~23%
- However, can not assume that all patients have this probability of readmission

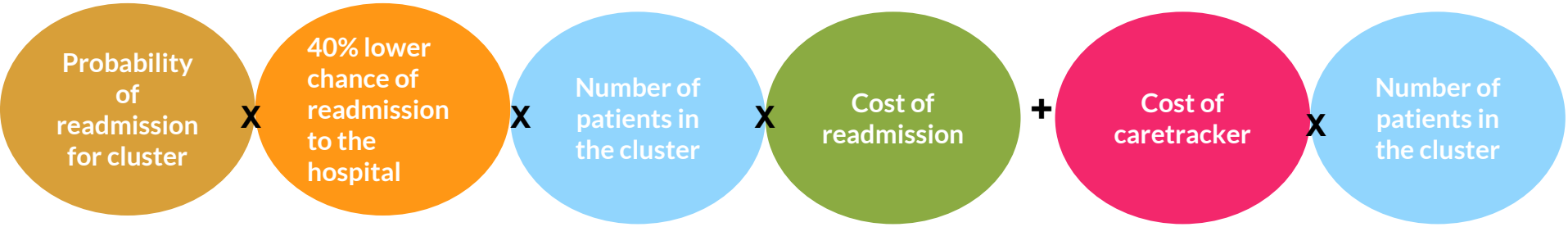


- Decision rule for clusters with higher readmission probability

Cost decision rule

Compare each to the cheapest baseline cost

Assign to Caretracker (expected cost)



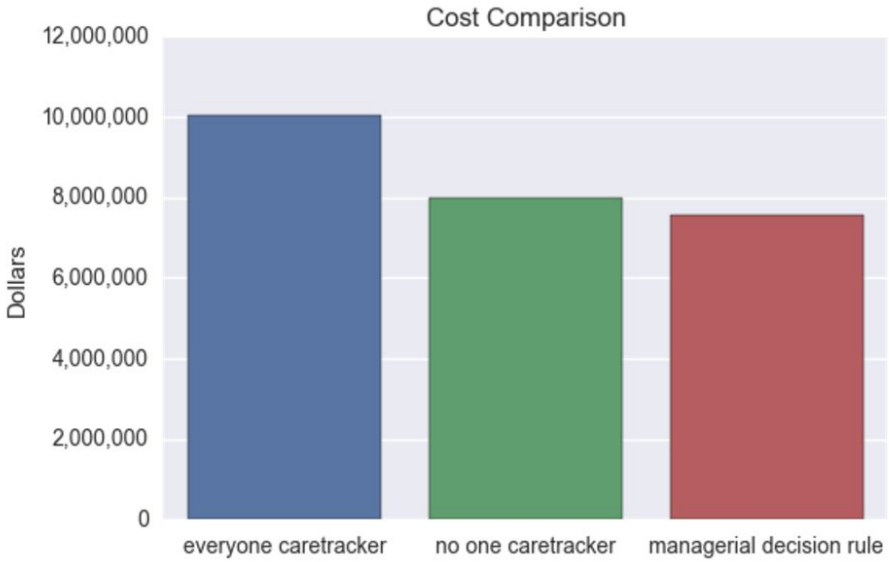
VS

Do not assign to Caretracker (expected cost)

- Cost of readmission is $8,000 per patient
- Cost of caretracker is $1,200 per patient

Managerial Clusters' Impact

- Assign caretracker to clusters with readmit probability >= 40%
- 850 people assigned to caretracker (6 clusters)
- **Savings: ~$430K**



Managerial clustering overview

- Easy to implement
- Assumes each cluster readmit probability is stationary over time
- Not individualized per patient
- Only considers three factors

- *Highest probability clusters contained comorbidity score over 131 and severity score over 19*
- *Note: The managerial decision rule cost is an expected cost, while the baseline (caretrack and no caretracker) is an actual cost*

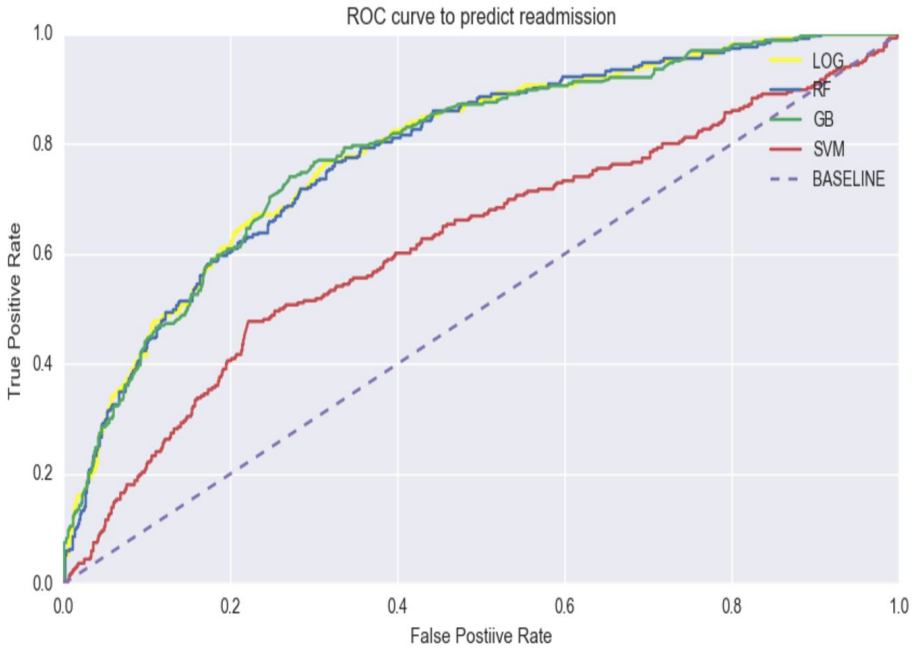
Machine Learning

Predict individualized readmission rate probability
Goal: Minimize annual cost

Classification Models Overview

- Default Accuracy
  - 77%
- Random Forest
  - Accuracy: 78%
- Gradient Boosting
  - Accuracy: 78.8%
- Logistic Classification
  - Accuracy: 79.5%
- Support Vector Machines
  - Accuracy: 76%

- Also, compare true positives, false positives, true negatives, and false negatives between models

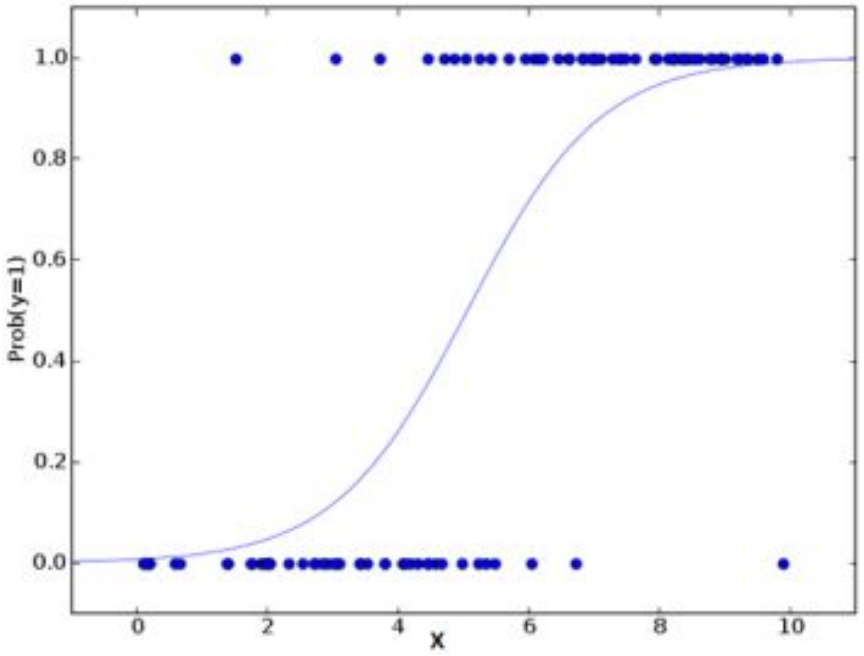


- *Default accuracy assumes no one is readmitted*
- *All models trained using three fold cross-validation*

Logistic Classification

- Similar to linear regression, except predict a probability of each patient being readmitted or not

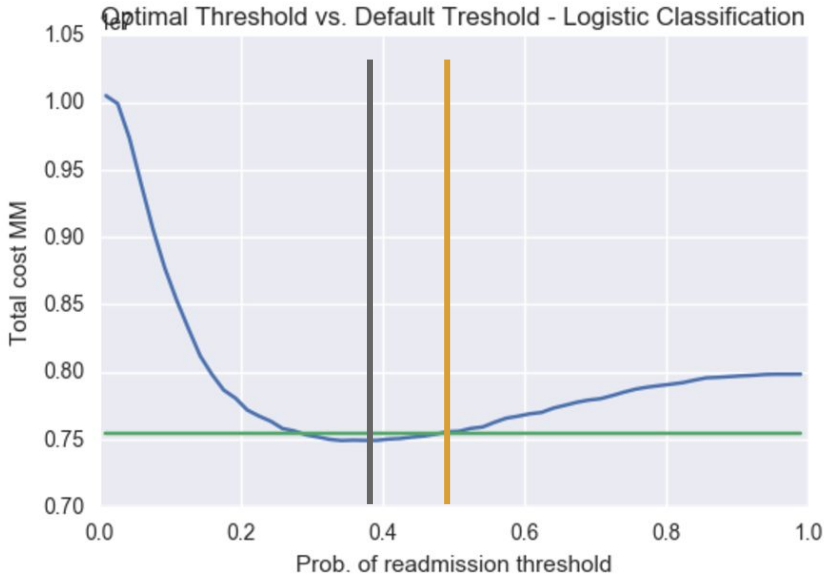
- By default, probability over 50% indicates that a person will be readmitted (threshold)



Optimize Logistic Classification

Parameters

- Same decision cost function from the managerial clusters
- Use the probability of readmission from logistic model to optimize cost
  - Ex: If a person in a cluster has a low probability of readmission, no need to assign to caretracker
- Goal: find the threshold of predicted probability (from logistic model) over which to assign people to caretracker



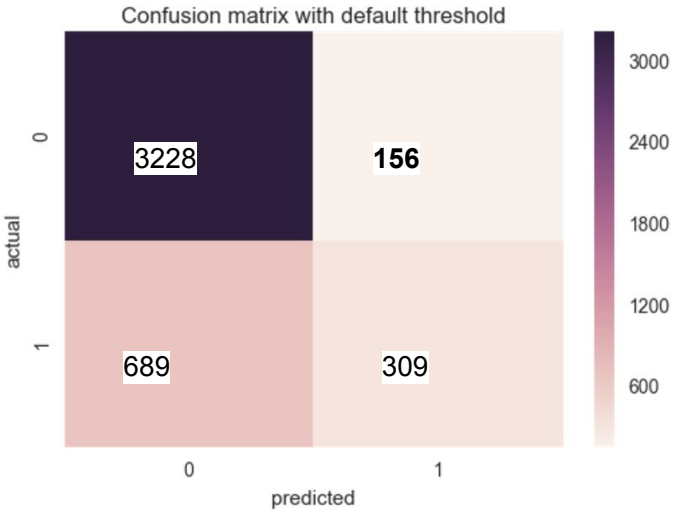
Default (.5) threshold

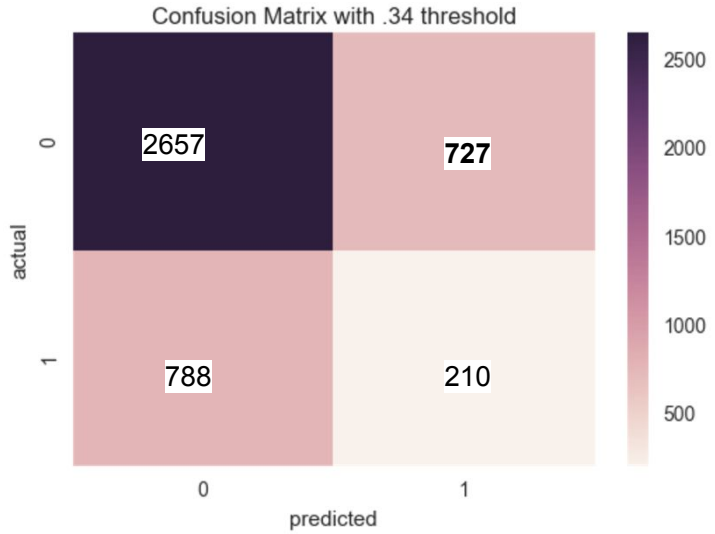
Optimal (.34) threshold

* Savings of **~$50k** between thresholds

Results From Logistic Classification

Patients per predicted category





- Predicted 1, actual 1 = true positive
  - Cost: $1.2K*Num. Patients + $8K *.6*Num. Patients
- Predicted 0, actual 1 = false negative
  - Cost: $8k * Num. Patients
- Predicted 1, actual 0 = false positive
  - Cost: $1.2K * Num. Patients
- Predicted 0, actual 0 = true negative
  - Cost: $0

Logistic Classification Overview

- Model is individualized per patient
  - Considers all six features
- The model needs to be re-run monthly as new readmission data comes in
- RMSE of .45 indicates high variability of predicting patient readmission
- Harder to implement as a production system compared to managerial clusters

How much can Tahoe Healthcare save?

$7.9 MM

Cheapest baseline (no one assigned to caretracker)

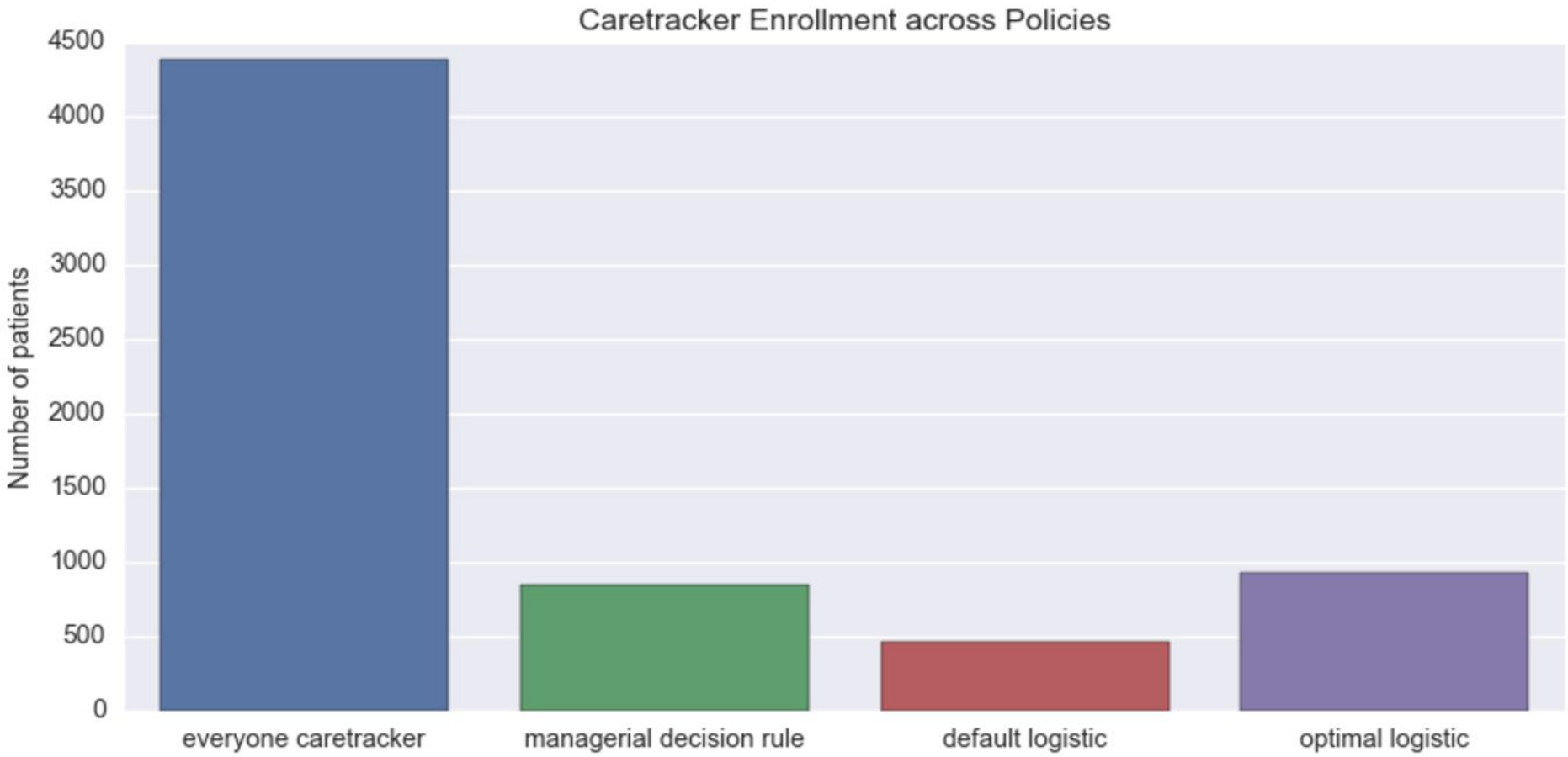
$7.49MM

Expected cost optimal logistic

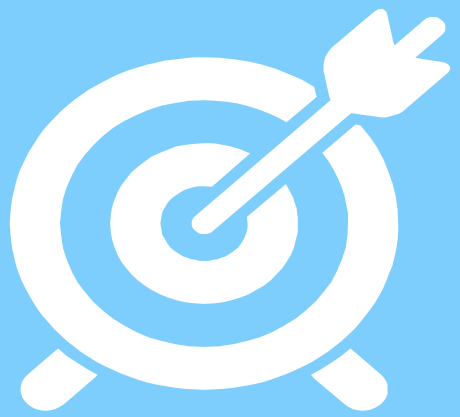
~$.49MM

Total expected savings for current patient database

Patients Enrolled In Caretracker



- Additional patients assigned to caretracker under optimal logistic is an indication of the effectiveness of individualization



Recommendations

Conclusion

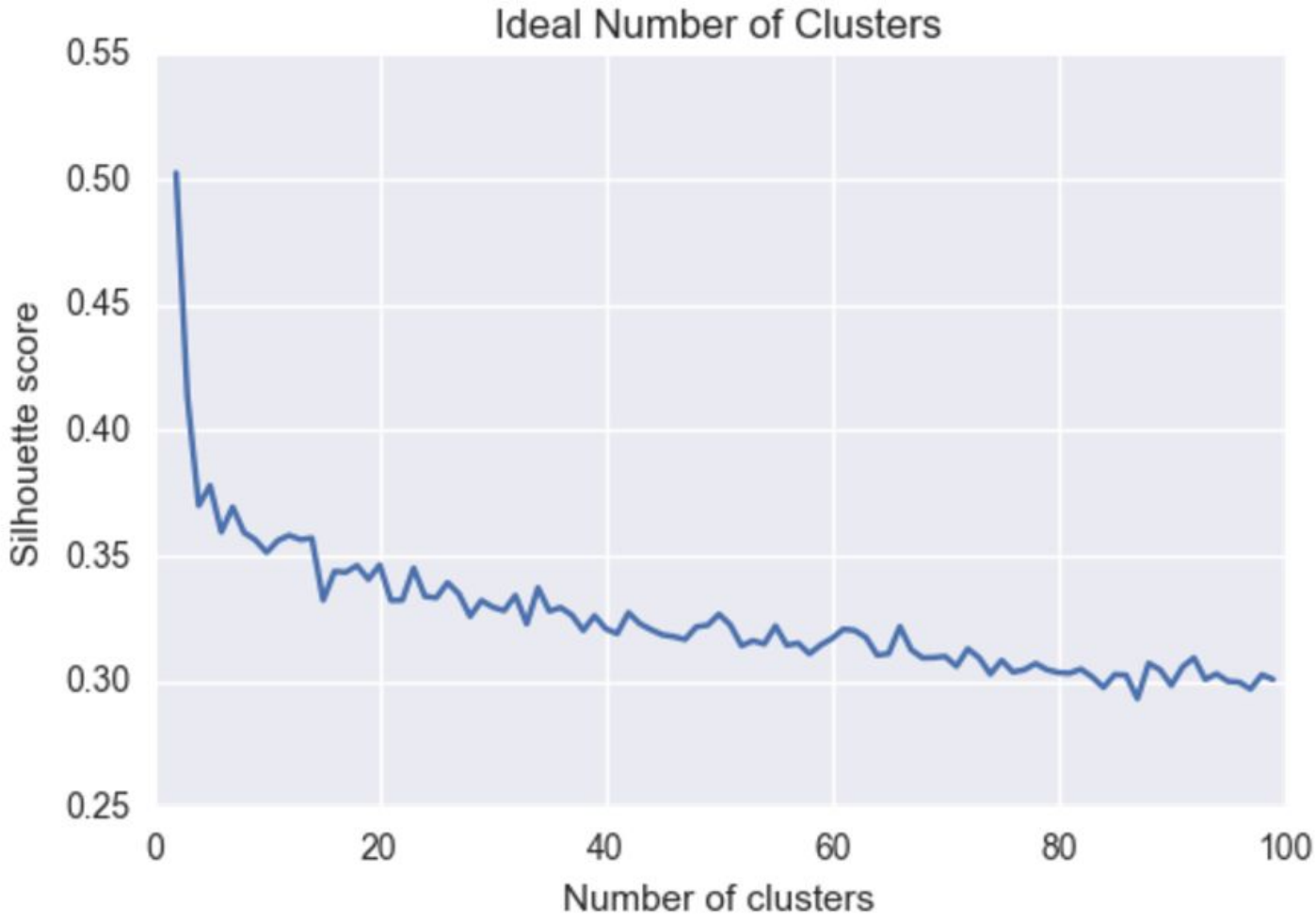
- Implement managerial clustering for savings of **$430k**
  - Even though logistic will save an additional **$57k**
- Lower maintenance and easier to implement
- No need to hire Data Scientist
  - Cost of hire **$100k** > savings of **$57k**
- Next steps
  - 1) Access to additional patients to validate cluster probabilities
  - 2) Additional features per patient
  - 3) Explore reducing caretracker cost

Questions

Business clusters assigned to caretracker

- 31: comorbidity_score above 131, severity_score above 32, in flu season
- 29: comorbidity_score above 131, severity_score 19-32, in flu season
- 25: comorbidity_score above 131, severity_score below 8, in flu season
- 27: comorbidity_score above 131, severity_score 8 to 19, in flu season
- 30: comorbidity_score above 131, severity_score above 32, not in flu season
- 28: comorbidity_score above 131, severity_score 19-32, not in flu season

Statistical patient clustering



Cost Comparison

The total ACTUAL cost of assigning everyone to caretracker is 10,048,800.00
The total ACTUAL cost of assigning noone to caretracker 7,984,000.00
The total EXPECTED cost of using managerial clusters and assigning to caretracker from there is 7,557,600.00
The EXPECTED cost of using logistic with the default threshold 7,550,000.00
The EXPECTED cost of using logistic with the optimal (.34) threshold 7,492,400.00
The savings of using the optimal logistic model is 491,600.00

Let's review some concepts

Yellow

Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.

Blue

Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.

Red

Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.

Yellow

Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.

Blue

Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.

Red

Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.

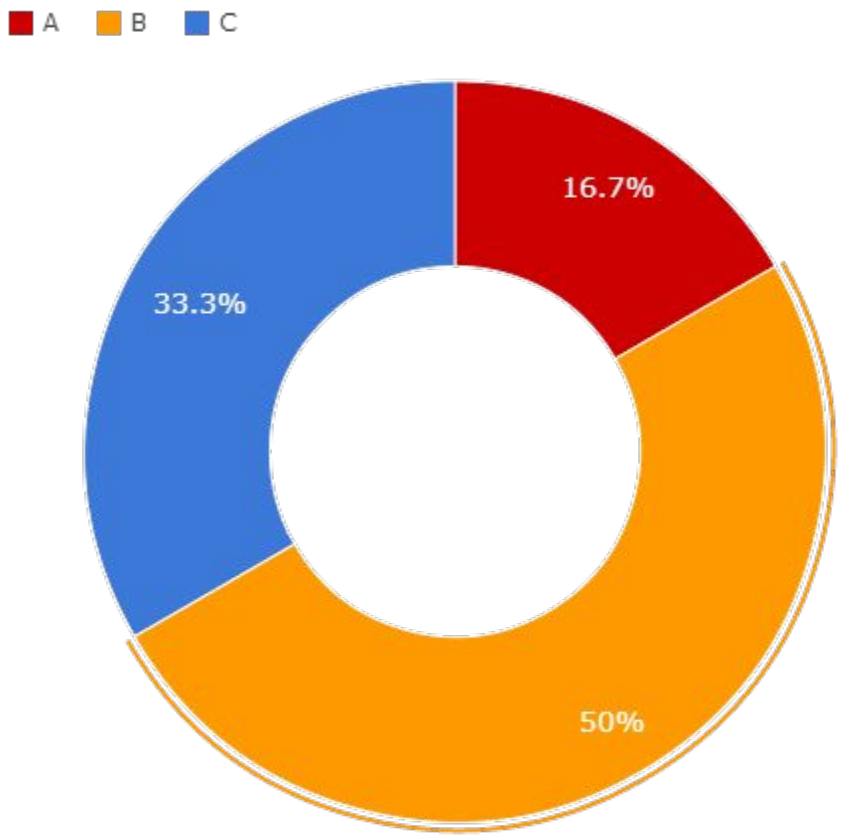
Thanks!

Any questions?

You can find me at:

@username

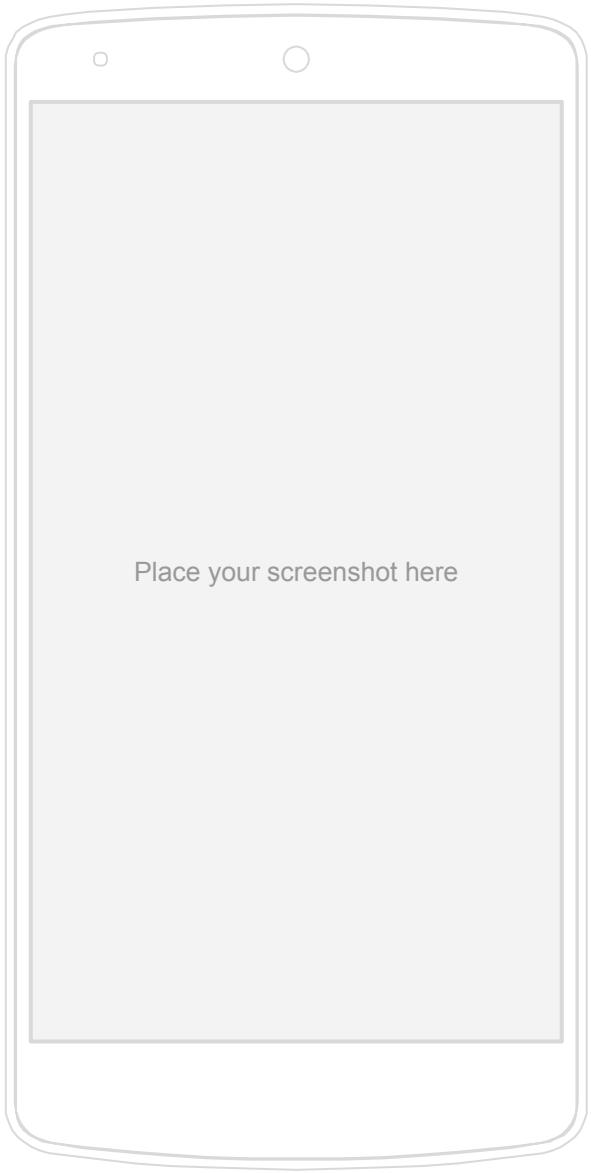
user@mail.me



You can copy&paste graphs from Google Sheets

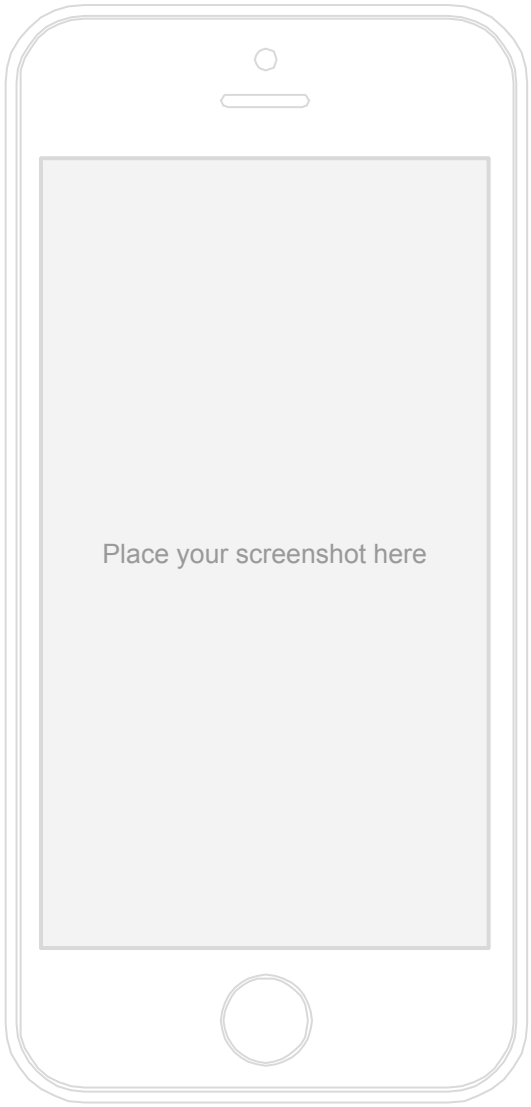
Android project

Show and explain your web, app or software projects using these gadget templates.



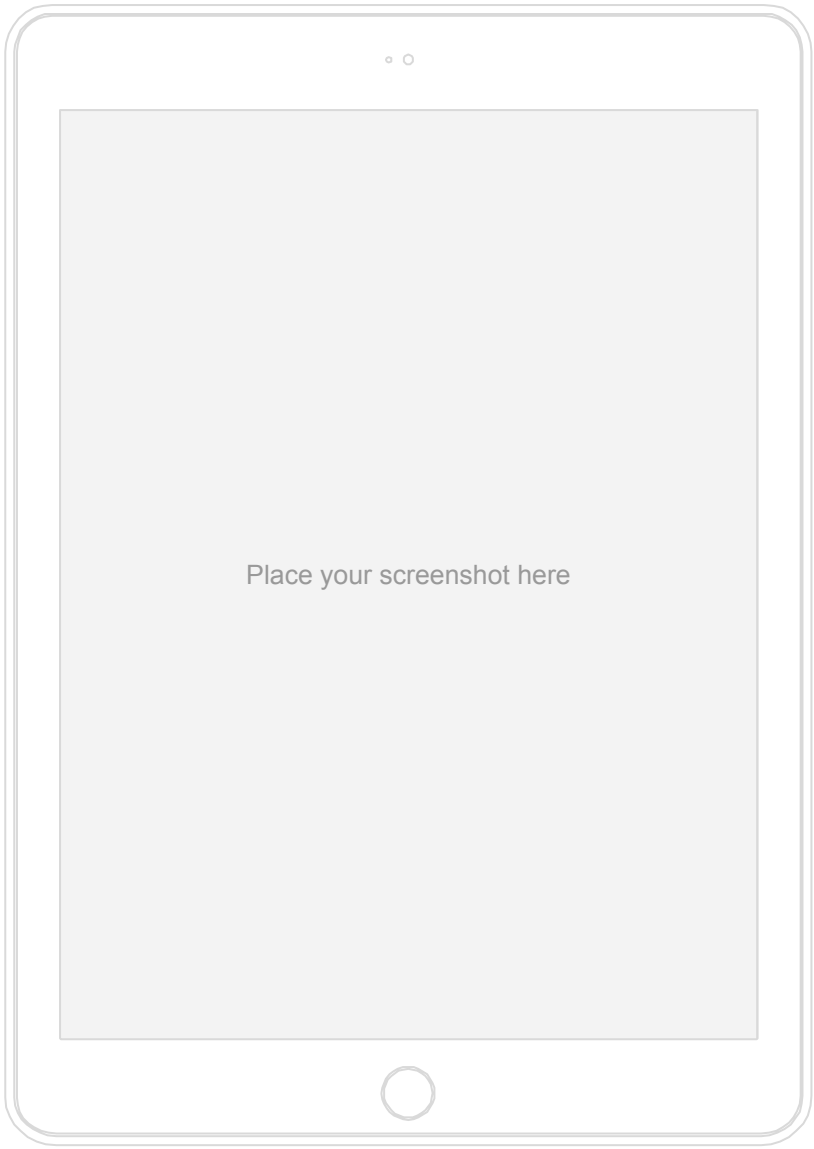
iPhone project

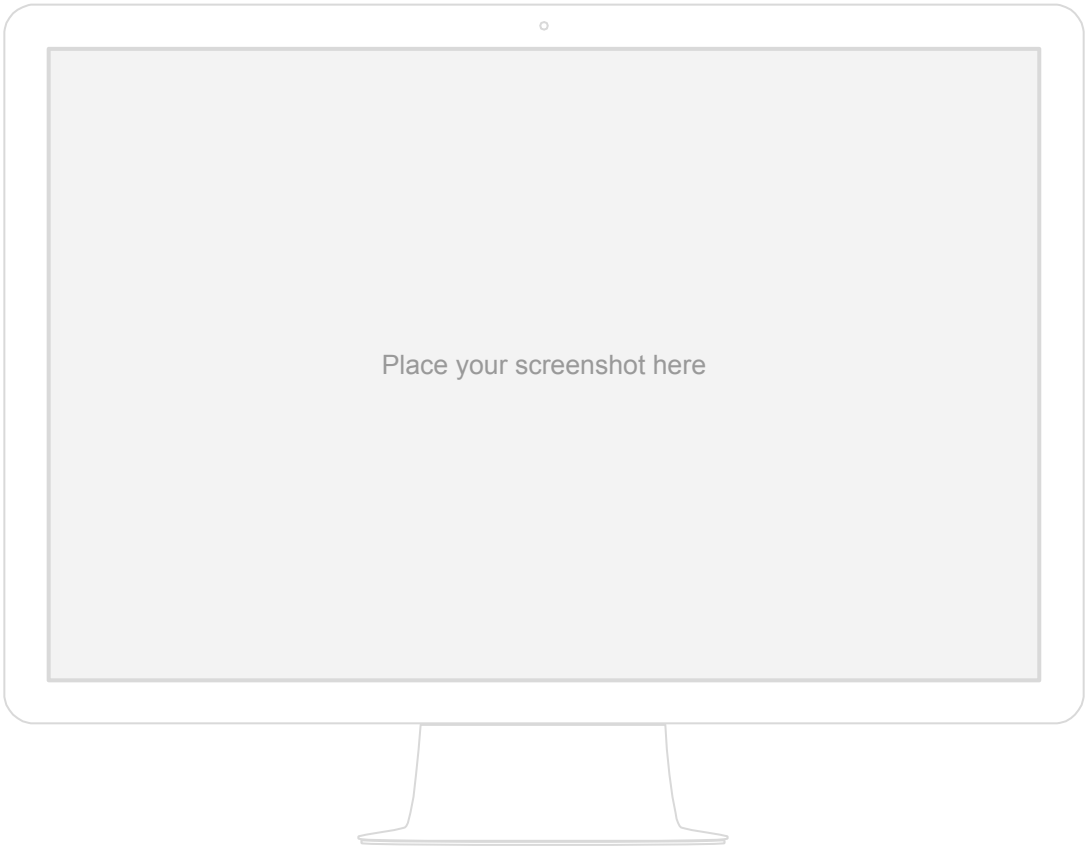
Show and explain your web, app or software projects using these gadget templates.



Tablet project

Show and explain your web, app or software projects using these gadget templates.





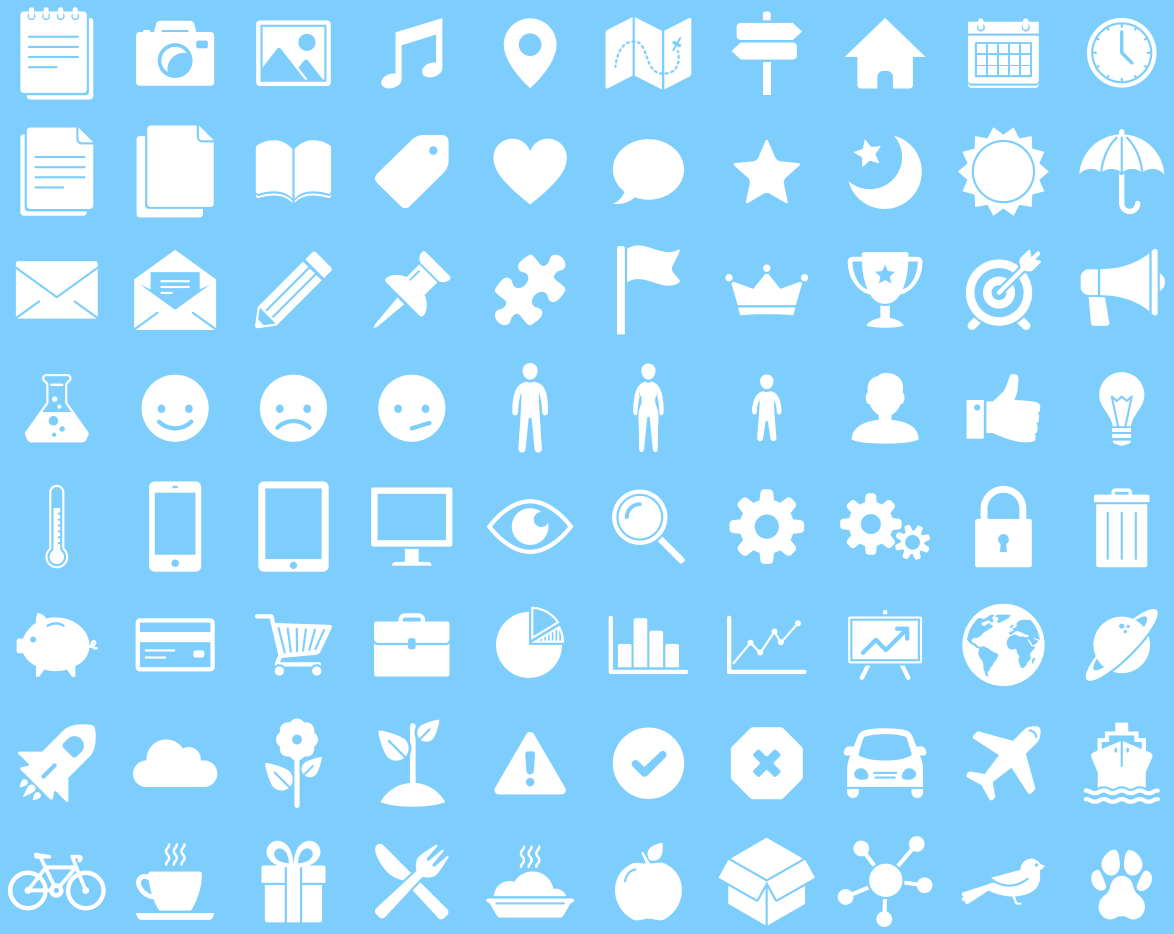
Desktop project

Show and explain your web, app or software projects using these gadget templates.

Credits

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by SlidesCarnival
- Photographs by Unsplash



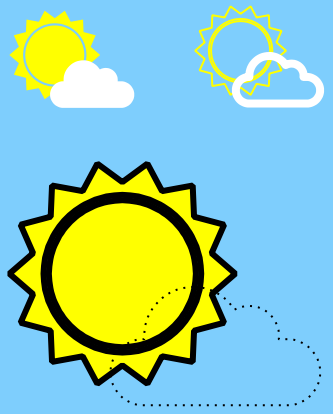
SlidesCarnival icons are editable shapes.

This means that you can:

- Resize them without losing quality.
- Change fill color and opacity.
- Change line color, width and style.

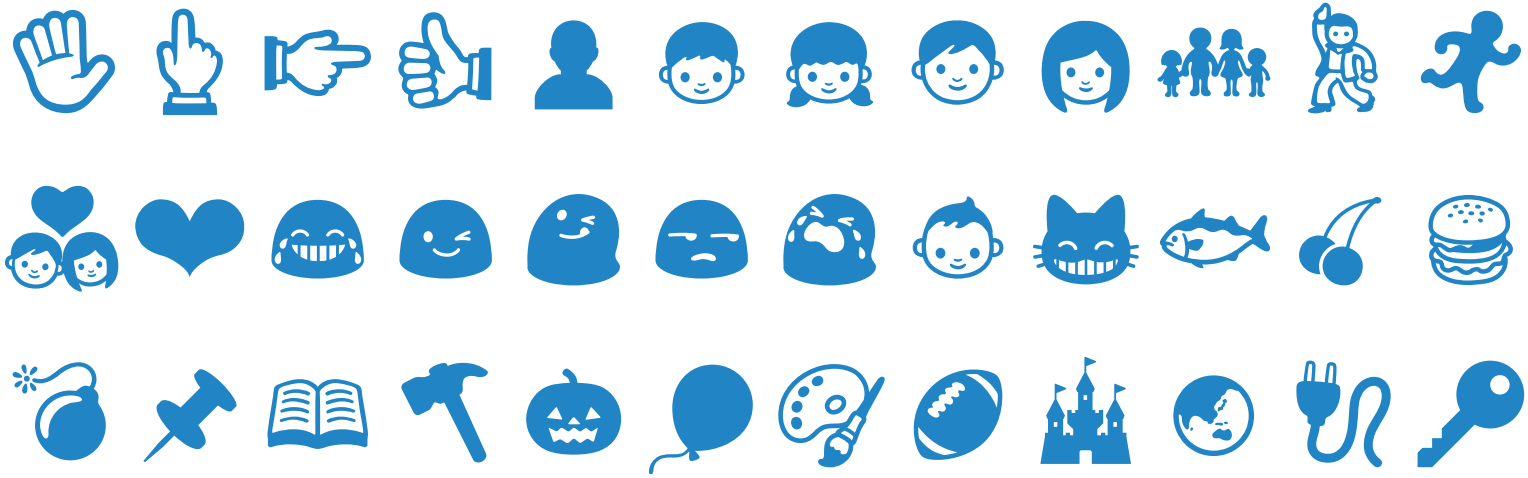
Isn't that nice? :)

Examples:



Now you can use any emoji as an icon!
And of course it resizes without losing quality and you can change the color.

How? Follow Google instructions
https://twitter.com/googledocs/status/730087240156643328



and many more...