Gamut

A Design Probe to Understand How Data Scientists Understand Machine Learning Models

CHI 2019

Fred Hohman
@fredhohman
Georgia Tech

Andrew Head
UC Berkeley

Rich Caruana
Microsoft Research

Rob DeLine
Microsoft Research

Steven Drucker
Microsoft Research

ai is

ai is **dangerous**
ai is **bad**
ai is **the new electricity**
ai is **good**
ai is **the future**
ai is **a crapshoot**
ai is **overhyped**
ai is **taking over**
ai is **coming**
ai is **scary**

Google Search | I'm Feeling Lucky

Report inappropriate predictions

*While **building and deploying** ML models is now an increasingly common practice, **interpreting** models is not.*

What is interpretability?

What is interpretability?

Human understanding

of a system's…

What is interpretability?

Human understanding
of a system's…

- internals
  e.g., components
  [Gilpin, 2018]
- operations
  e.g., math
  [Biran, 2017]
- data mapping
  e.g., input to output
  [Montavon, 2017]
- representation
  in an explanation
  [Ribeiro, 2016]

What is interpretability?

Human understanding
of a system's…

- **internals**
  e.g., components
  [Gilpin, 2018]
- **operations**
  e.g., math
  [Biran, 2017]
- **data mapping**
  e.g., input to output
  [Montavon, 2017]
- **representation**
  in an explanation
  [Ribeiro, 2016]

No formal, agreed upon definition
[Lipton, 2016]

GDPR (General Data Protection Regulation)

GDPR (General Data Protection Regulation)

↳ Chapter 3 → Section 4 → Article 22

"Automated individual decision-making, including profiling"

GDPR (General Data Protection Regulation)

↳ Chapter 3 → Section 4 → Article 22

“Automated individual decision-making, including profiling”

⚖️ *Right to explanation*

Gamut Contributions

1. **Capabilities**
   of interpretability
2. **Design Probe**
   embodying capabilities
3. **Evaluation & Investigation**
   of probe & emerging practice of interpretability w/ real users

Contribution 1: Interpretability Capabilities

Can we operationalize interpretability?

Can we operationalize interpretability?

Formative research with professional data scientists @

- 4 senior ML researchers
- 5 ML practitioners

Can we operationalize interpretability?

Formative research with professional data scientists @

- 4 senior ML researchers
- 5 ML practitioners

Prompt: *In a perfect world, given a machine learning model, what questions would you ask it to help you interpret both the model and its predictions?*

From formative research

Explainable ML Interface Questions

From formative research

Explainable ML Interface Questions

From formative research

Explainable ML Interface Questions

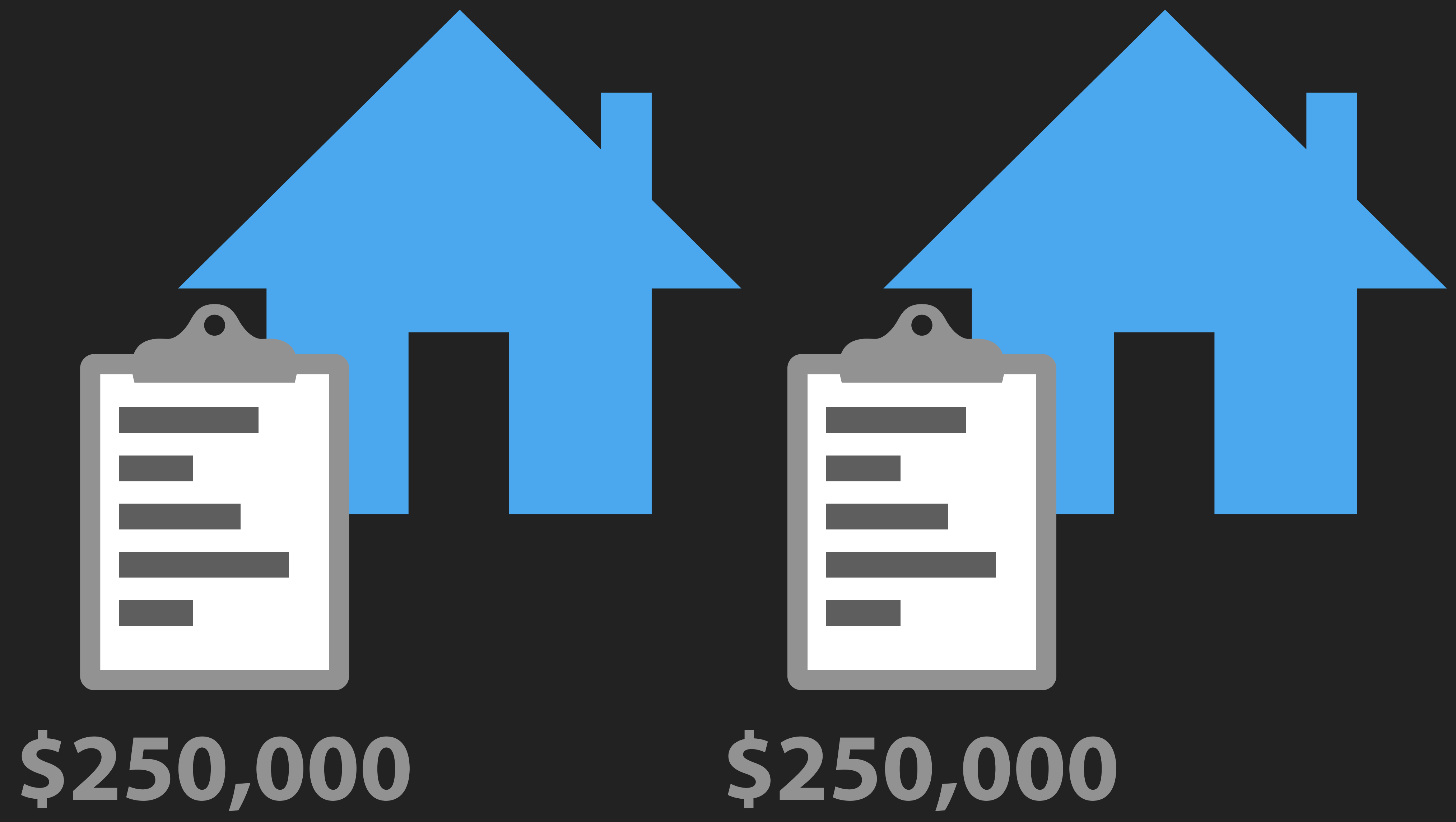
Why does this house cost that much?

From formative research

Explainable ML Interface Questions

Why does this house cost that much?

What is the difference between these two?

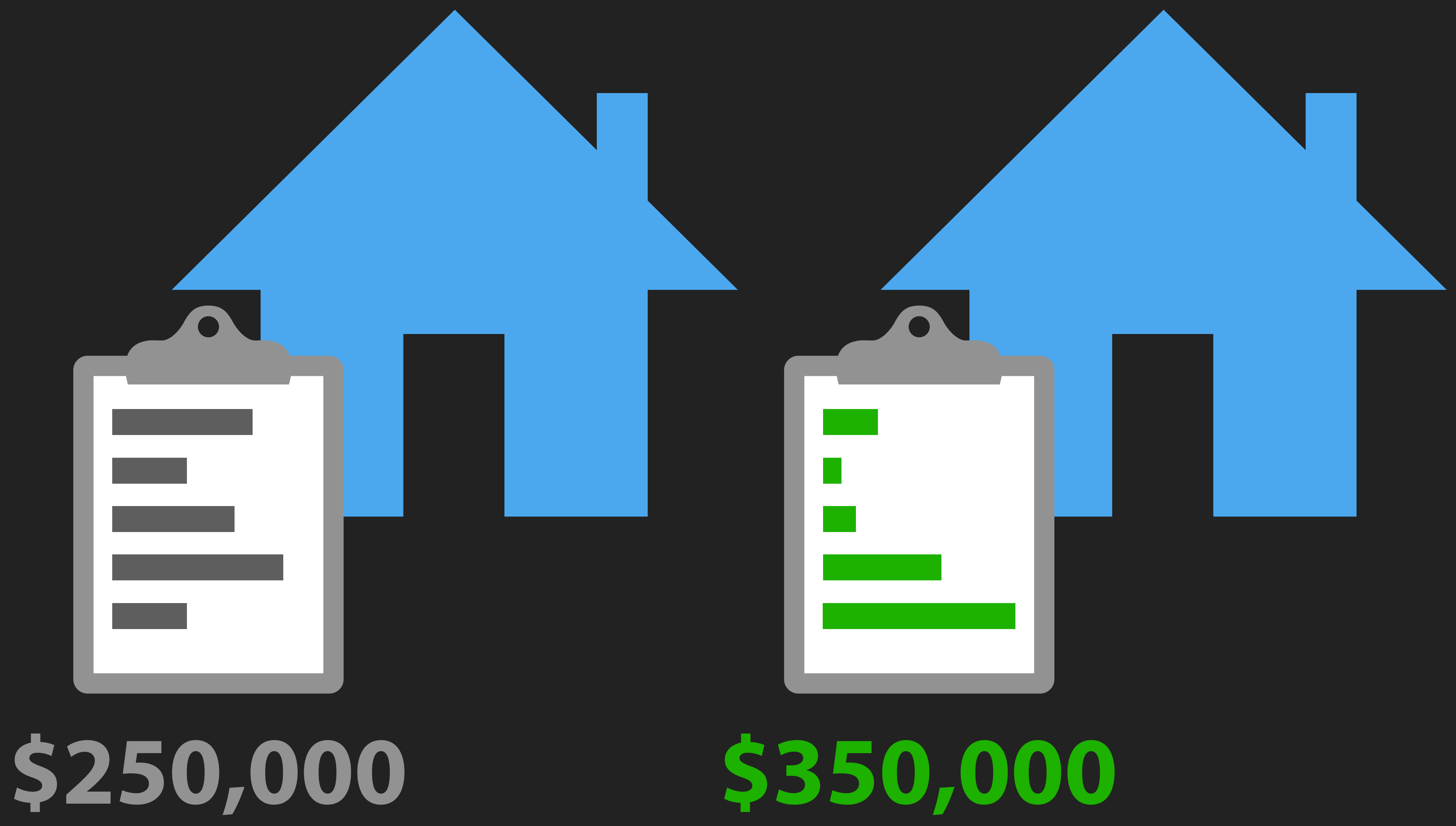


From formative research

Explainable ML Interface Questions

Why does this house cost that much?

What is the difference between these two?



From formative research

Explainable ML Interface Questions

Why does this house cost that much?

What is the difference between these two?

What if I added…

From formative research

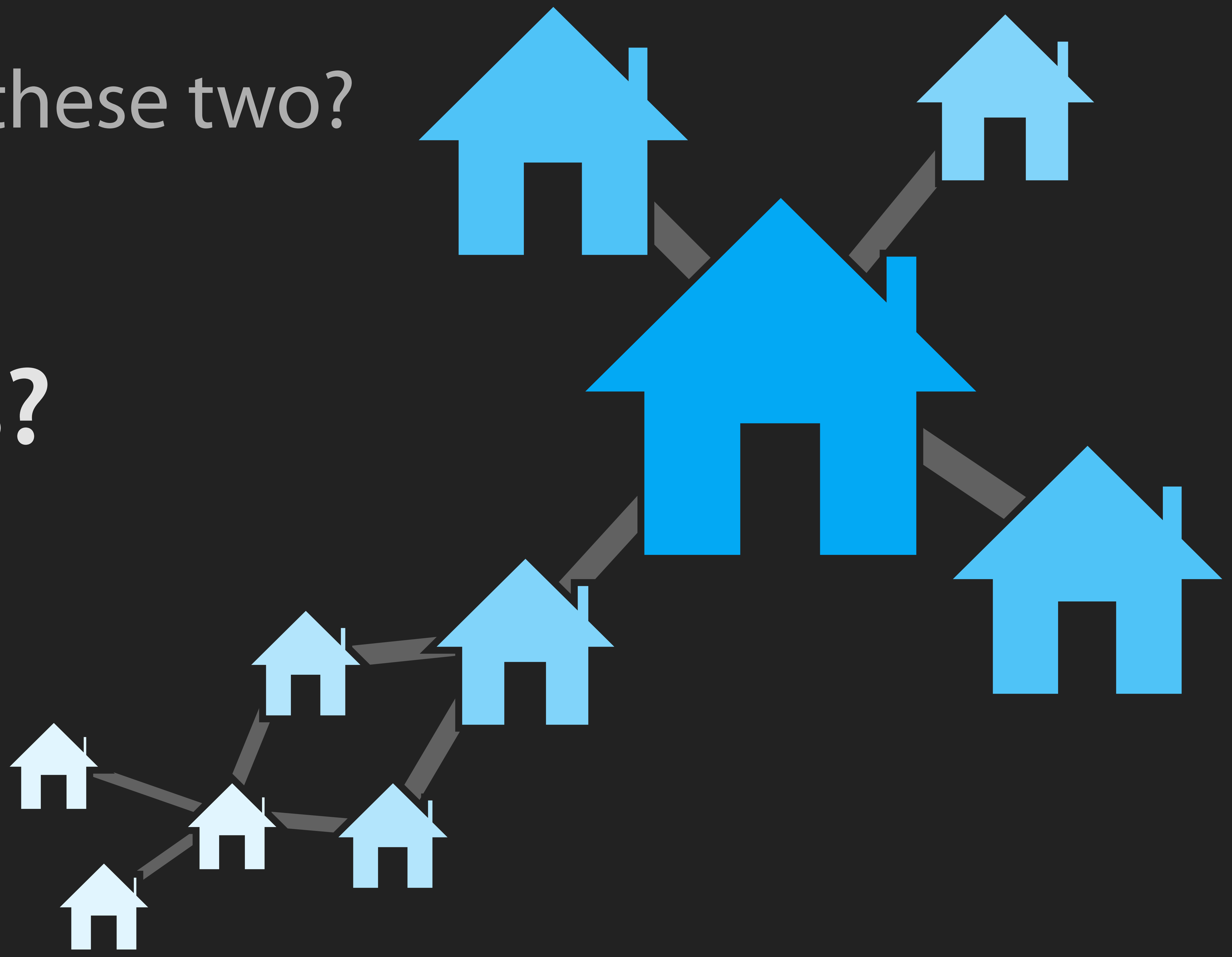
Explainable ML Interface Questions

Why does this house cost that much?

What is the difference between these two?

What if I added…

What are similar homes?



From formative research

Explainable ML Interface Questions

Why does this house cost that much?

What is the difference between these two?

What if I added…

What are similar homes?

Where is it wrong?

From formative research

Explainable ML Interface Questions

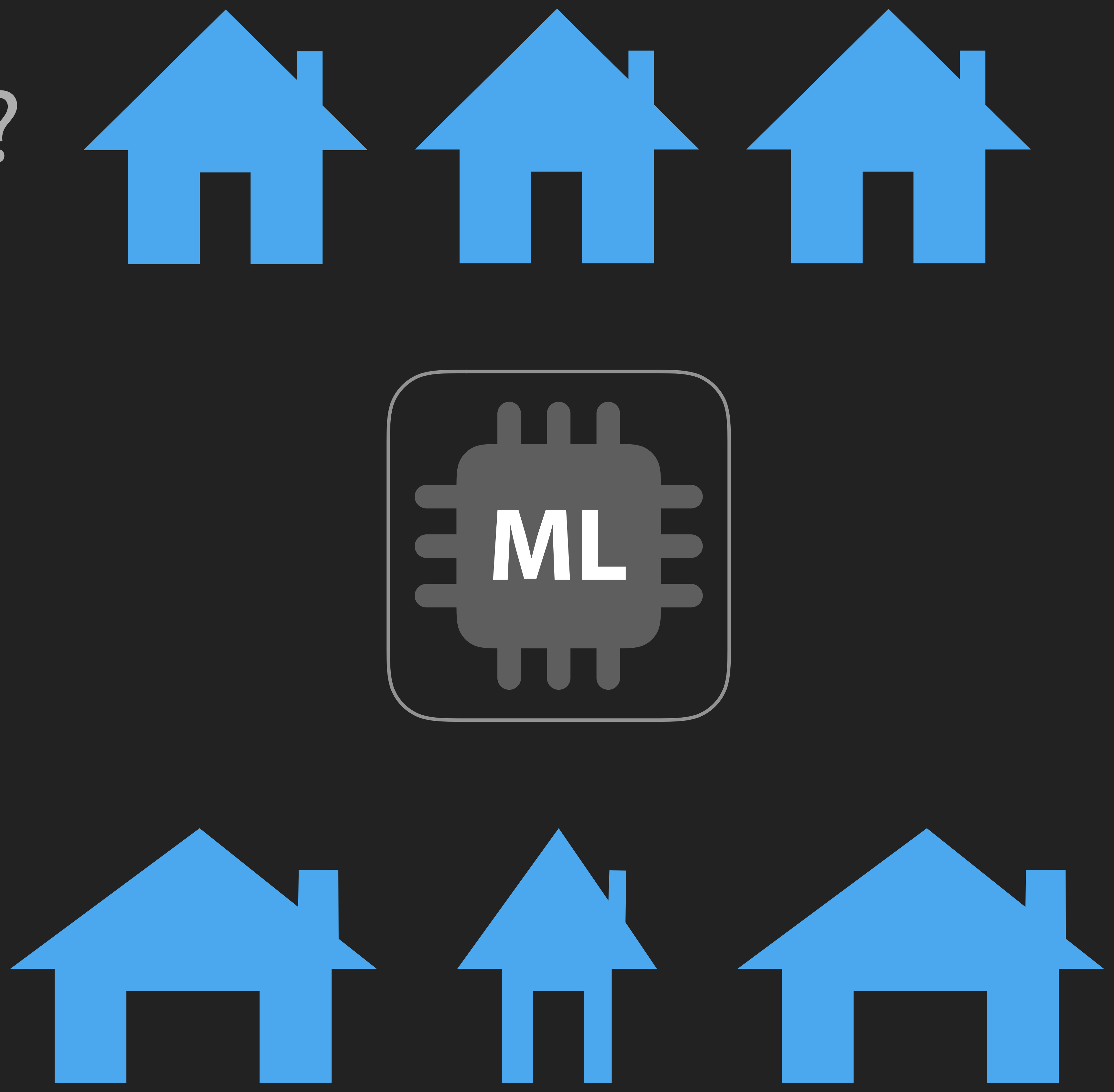
Why does this house cost that much?

What is the difference between these two?

What if I added…

What are similar homes?

Where is it wrong?



From formative research

Explainable ML Interface Questions

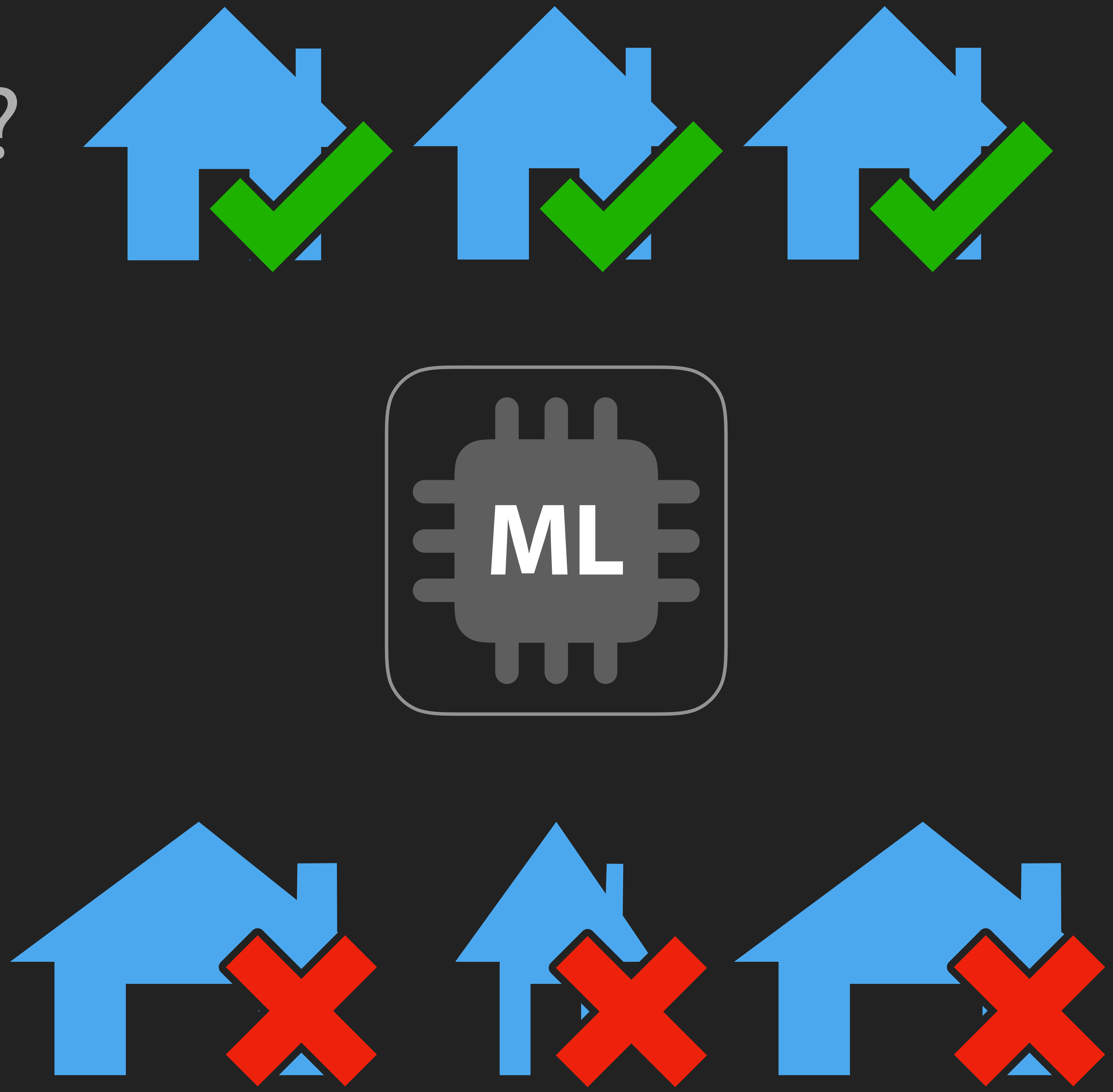
Why does this house cost that much?

What is the difference between these two?

What if I added…

What are similar homes?

Where is it wrong?



From formative research

Explainable ML Interface Questions

Why does this house cost that much?

What is the difference between these two?

What if I added…

What are similar homes?

Where is it wrong?

What is most important?

From formative research

Explainable ML Interface Questions

Why does this house cost that much?

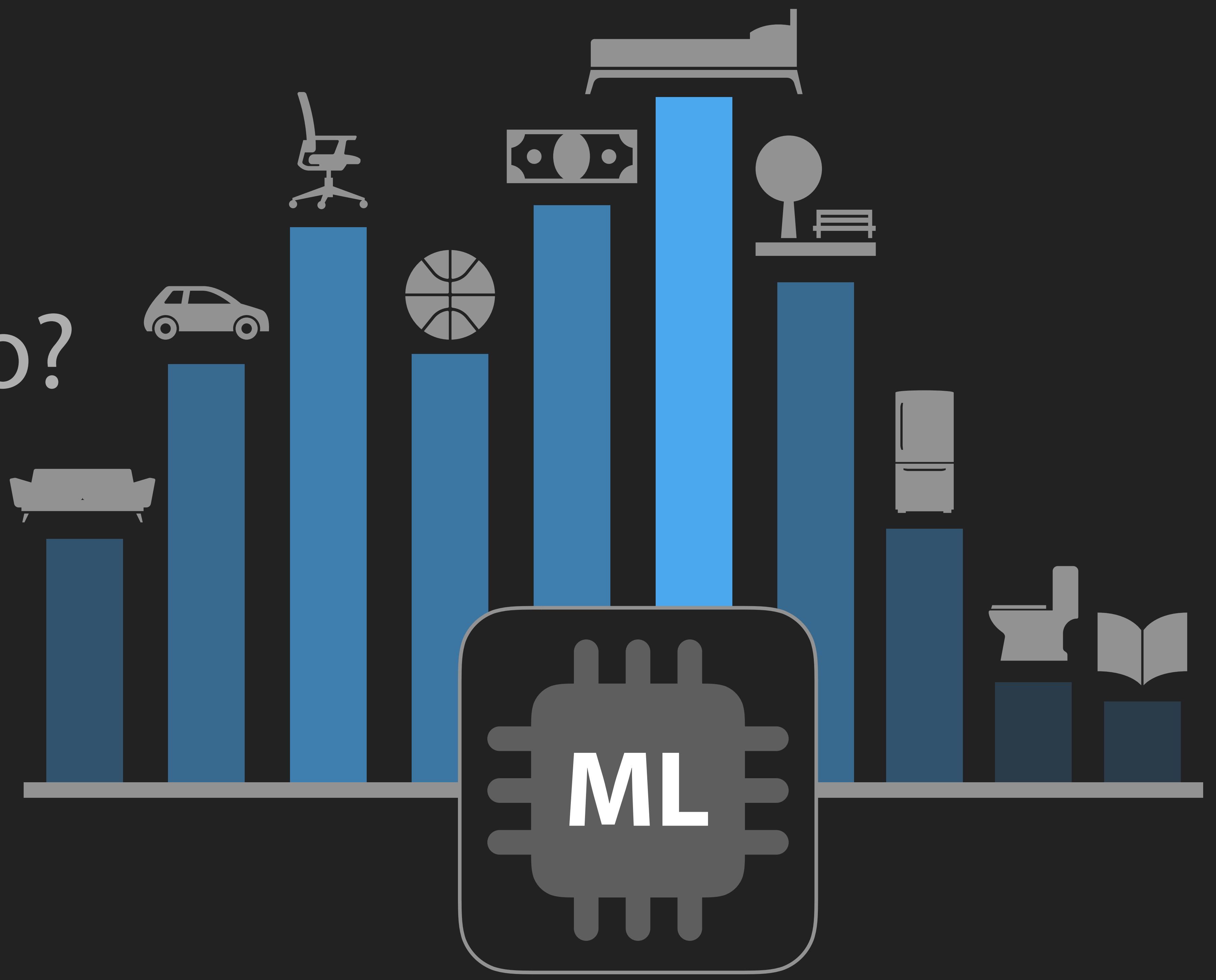
What is the difference between these two?

What if I added…

What are similar homes?

Where is it wrong?

What is most important?



From formative research

Explainable ML Interface Questions

Why does this house cost that much?

What is the difference between these two?

What if I added…

What are similar homes?

Where is it wrong?

What is most important?

From formative research

Explainable ML Interface Capabilities

C1 Why does this house cost that much?
Local instance explanations

C2 What is the difference between these two?
Instance explanation comparisons

C3 What if I added…
Counterfactuals

C4 What are similar homes?
Nearest neighbors

C5 Where is it wrong?
Regions of error

C6 What is most important?
Feature importance

From formative research

Explainable ML Interface **Capabilities**

C1 Why does this house cost that much?
Local instance explanations

C2 What is the difference between these two?
Instance explanation comparisons

C3 What if I added…
Counterfactuals

C4 What are similar homes?
Nearest neighbors

C5 Where is it wrong?
Regions of error

C6 What is most important?
Feature importance

GAMUT: A Design Probe to Understand How Data Scientists Understand Machine Learning Models

Fred Hohman
Georgia Institute of Technology
Atlanta, GA, USA
fredhohman@gatech.edu

Andrew Head
UC Berkeley
Berkeley, CA, USA
andrewhead@berkeley.edu

Rich Caruana
Microsoft Research
Redmond, WA, USA
rcaruana@microsoft.com

Robert DeLine
Microsoft Research
Redmond, WA, USA
rob.deline@microsoft.com

Steven M. Drucker
Microsoft Research
Redmond, WA, USA
sdrucker@microsoft.com

ABSTRACT

Without good models and the right tools to interpret them, data scientists risk making decisions based on hidden biases, spurious correlations, and false generalizations. This has led to a rallying cry for model interpretability. Yet the concept of interpretability remains nebulous, such that researchers and tool designers lack actionable guidelines for how to incorporate interpretability into models and accompanying tools. Through an iterative design process with expert machine learning researchers and practitioners, we designed a visual analytics system, GAMUT, to explore how interactive interfaces could better support model interpretation. Using GAMUT as a probe, we investigated why and how professional data scientists interpret models, and how interface affordances can support data scientists in answering questions about model interpretability. Our investigation showed that interpretability is not a monolithic concept: data scientists

CCS CONCEPTS

• **Human-centered computing → Empirical studies in visualization**; *Visualization systems and tools*; • **Computing methodologies → Machine learning**.

KEYWORDS

Machine learning interpretability, design probe, visual analytics, data visualization, interactive interfaces

ACM Reference Format:
Fred Hohman, Andrew Head, Rich Caruana, Robert DeLine, and Steven M. Drucker. 2019. GAMUT: A Design Probe to Understand How Data Scientists Understand Machine Learning Models. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3290605.3300809

Definitions + examples in the paper!

Contribution 2: Design Probe

How to test our capabilities?

Contribution 2: Design Probe

How to test our capabilities?

Goal: *understand emerging practice of model interpretability*

Contribution 2: Design Probe

How to test our capabilities?

Goal: *understand emerging practice of model interpretability*

[Hutchinson, 2003]

Design probe: *"instrument that is deployed to find out about the unknown—returning with useful or interesting data."*

Balance of ***design***, ***social science***, ***engineering***

How does our design probe support our capabilities?

House 550

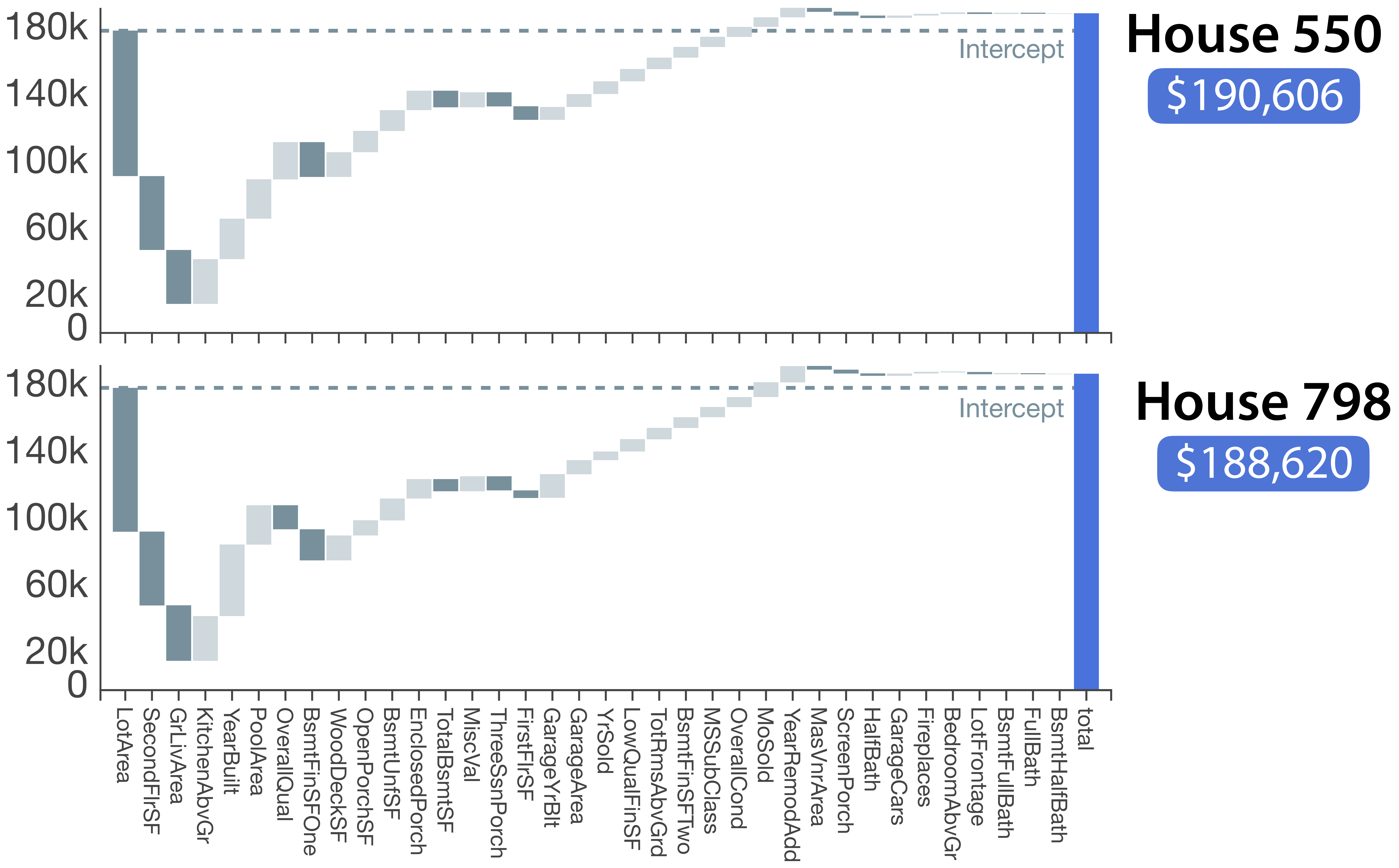
$190,606

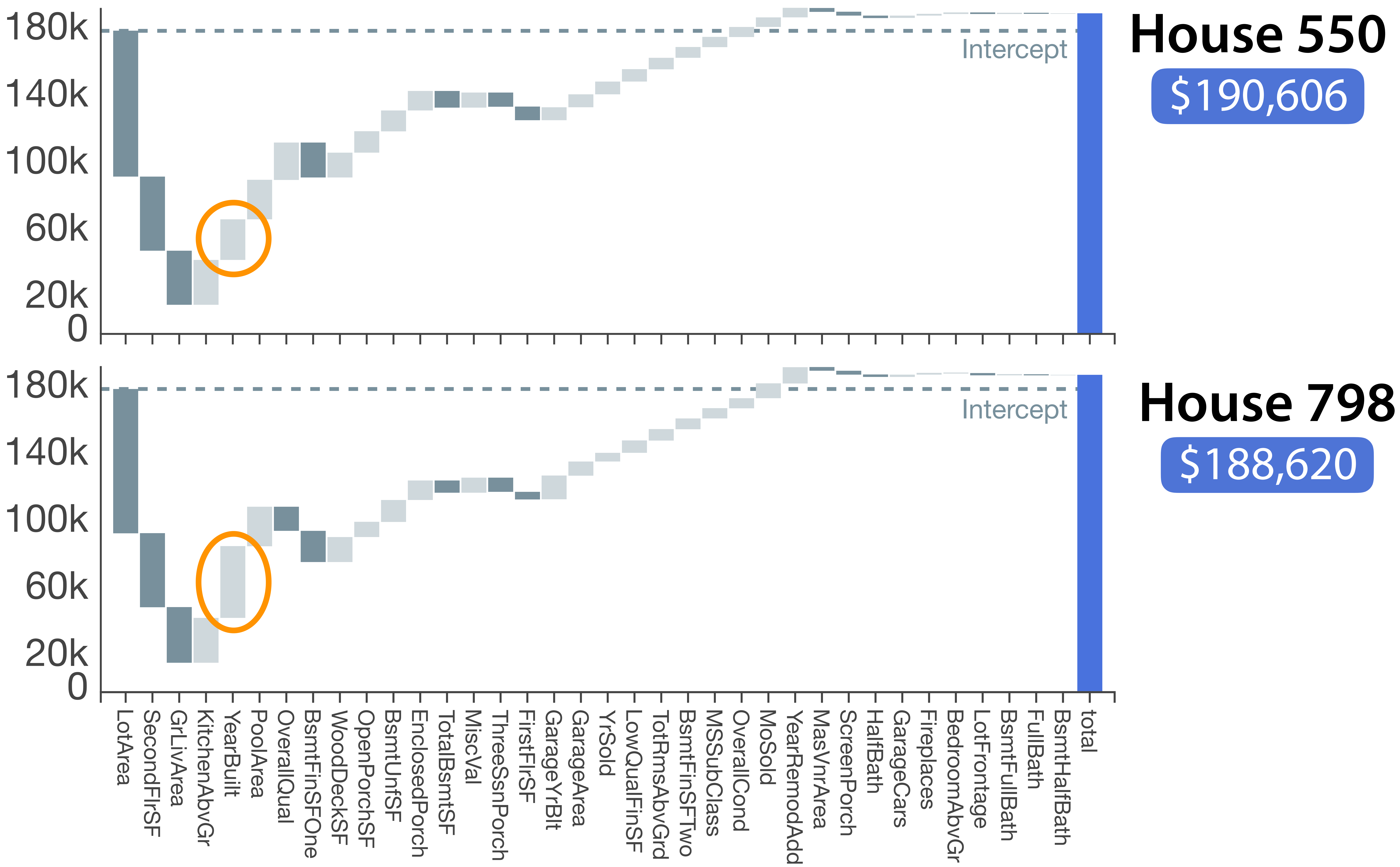
House 550

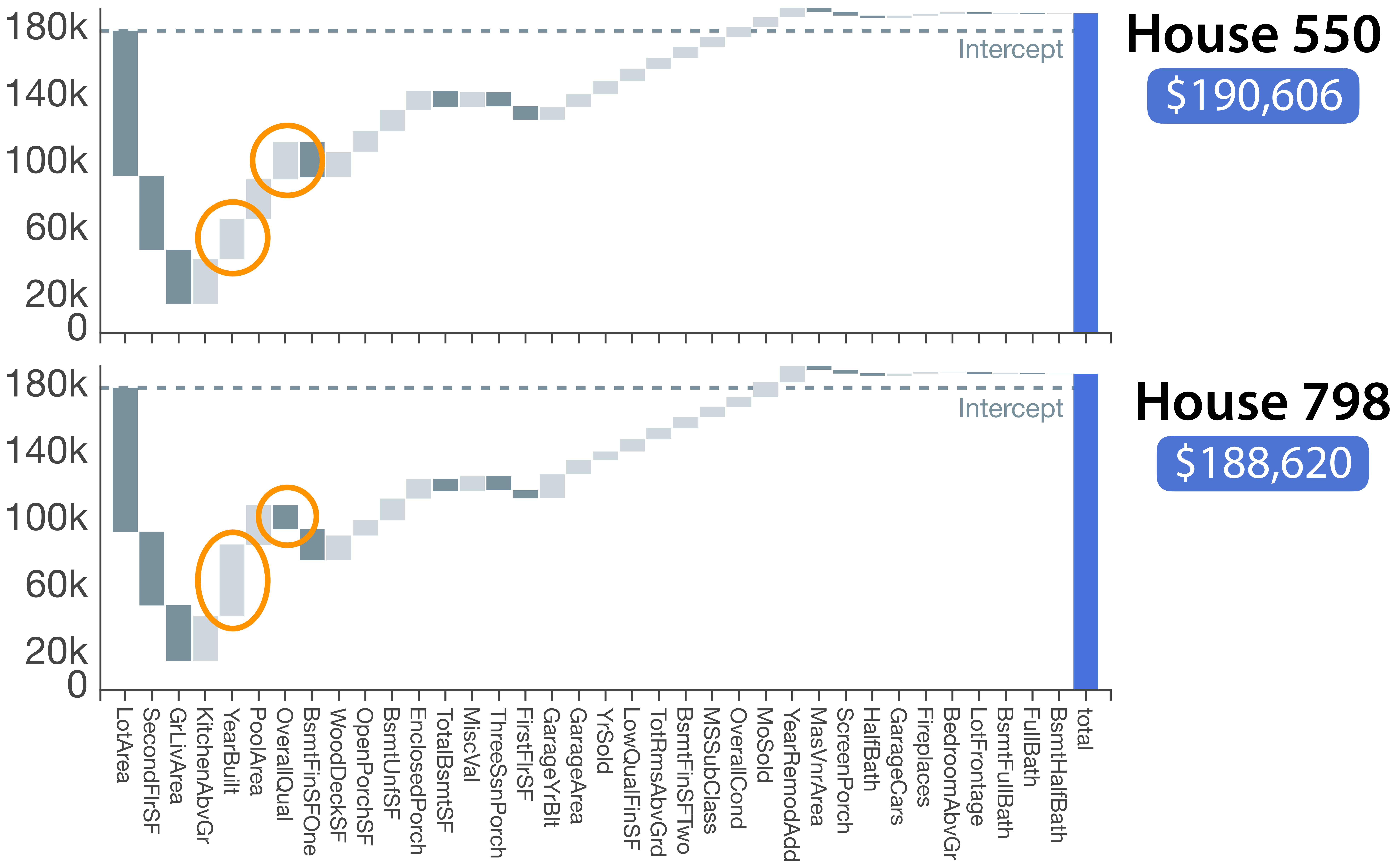
$190,606

BsmtHalfBath
FullBath
BsmtFullBath
LotFrontage
BedroomAbvGr
Fireplaces
GarageCars
HalfBath
ScreenPorch
MasVnrArea
YearRemodAdd
MoSold
OverallCond
MSSubClass
BsmtFinSFTwo
TotRmsAbvGrd
LowQualFinSF
YrSold
GarageArea
GarageYrBlt
FirstFlrSF
ThreeSsnPorch
MiscVal
TotalBsmtSF
EnclosedPorch
BsmtUnfSF
OpenPorchSF
WoodDeckSF
BsmtFinSFOne
OverallQual
PoolArea
YearBuilt
KitchenAbvGr
GrLivArea
SecondFlrSF
LotArea

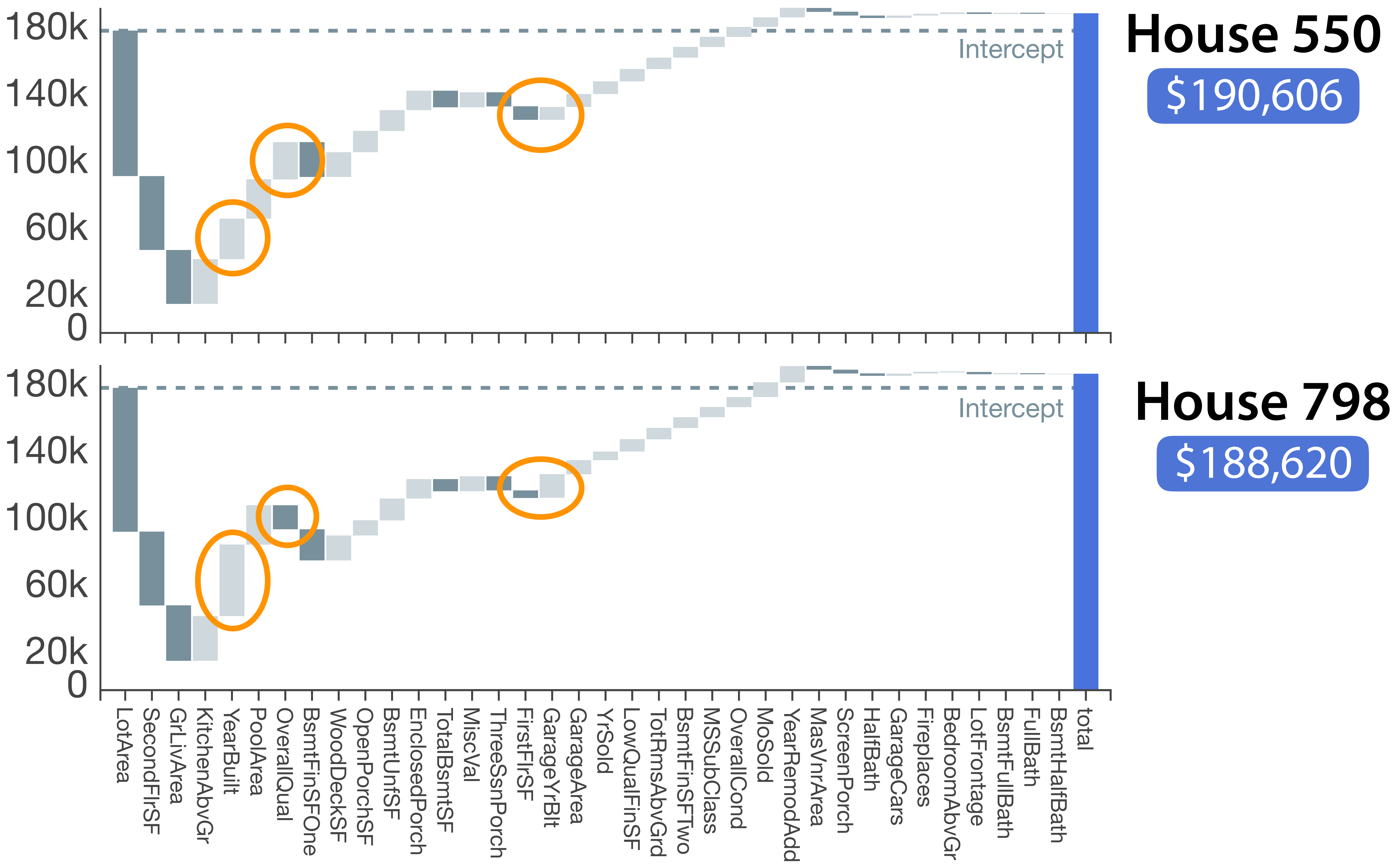

House 550
$190,606
Prediction
180k
140k
100k
60k
20k
0
Intercept
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total
House features

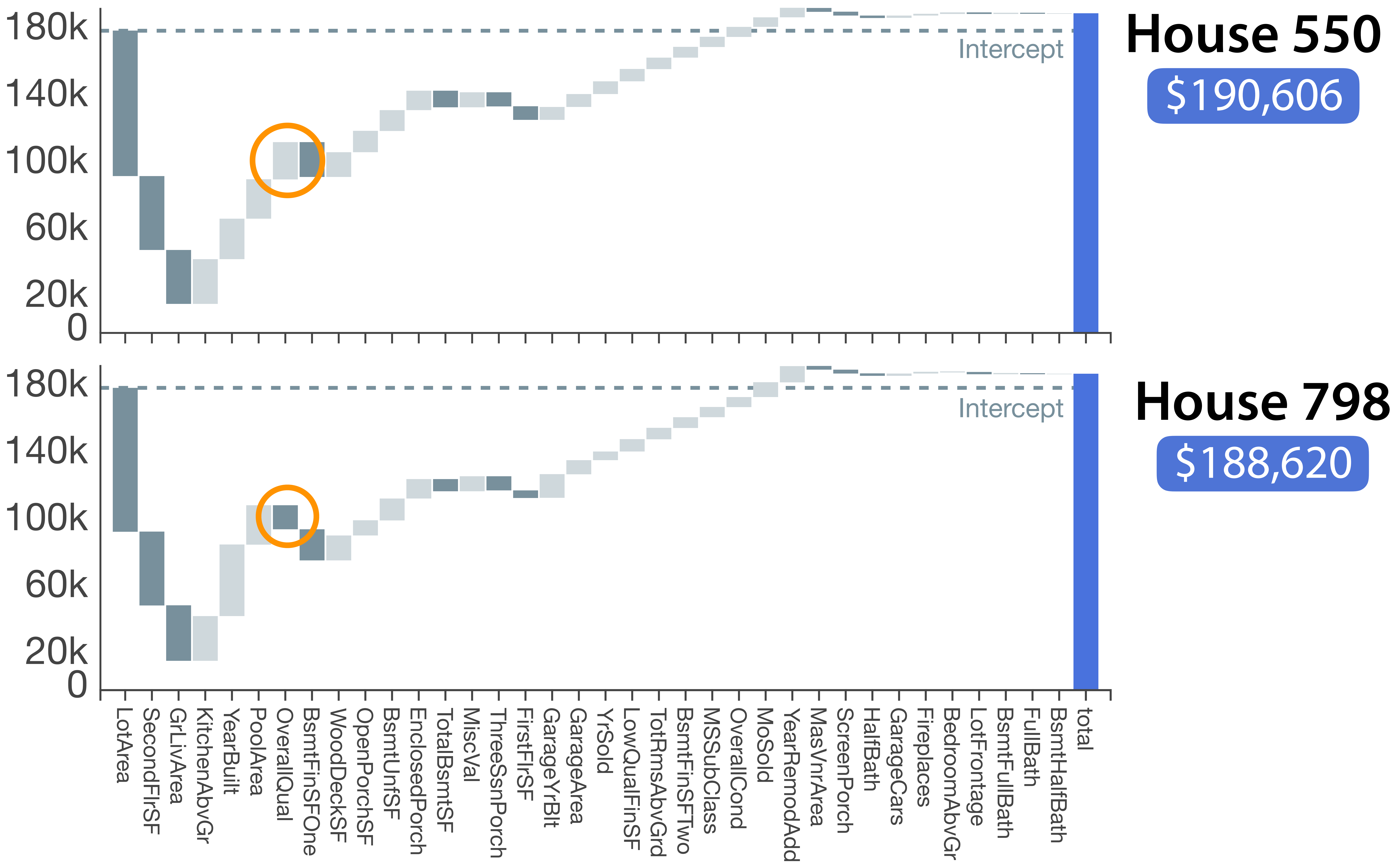

House 550
$190,606
House 798
$188,620
Intercept
180k
140k
100k
60k
20k
0
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total

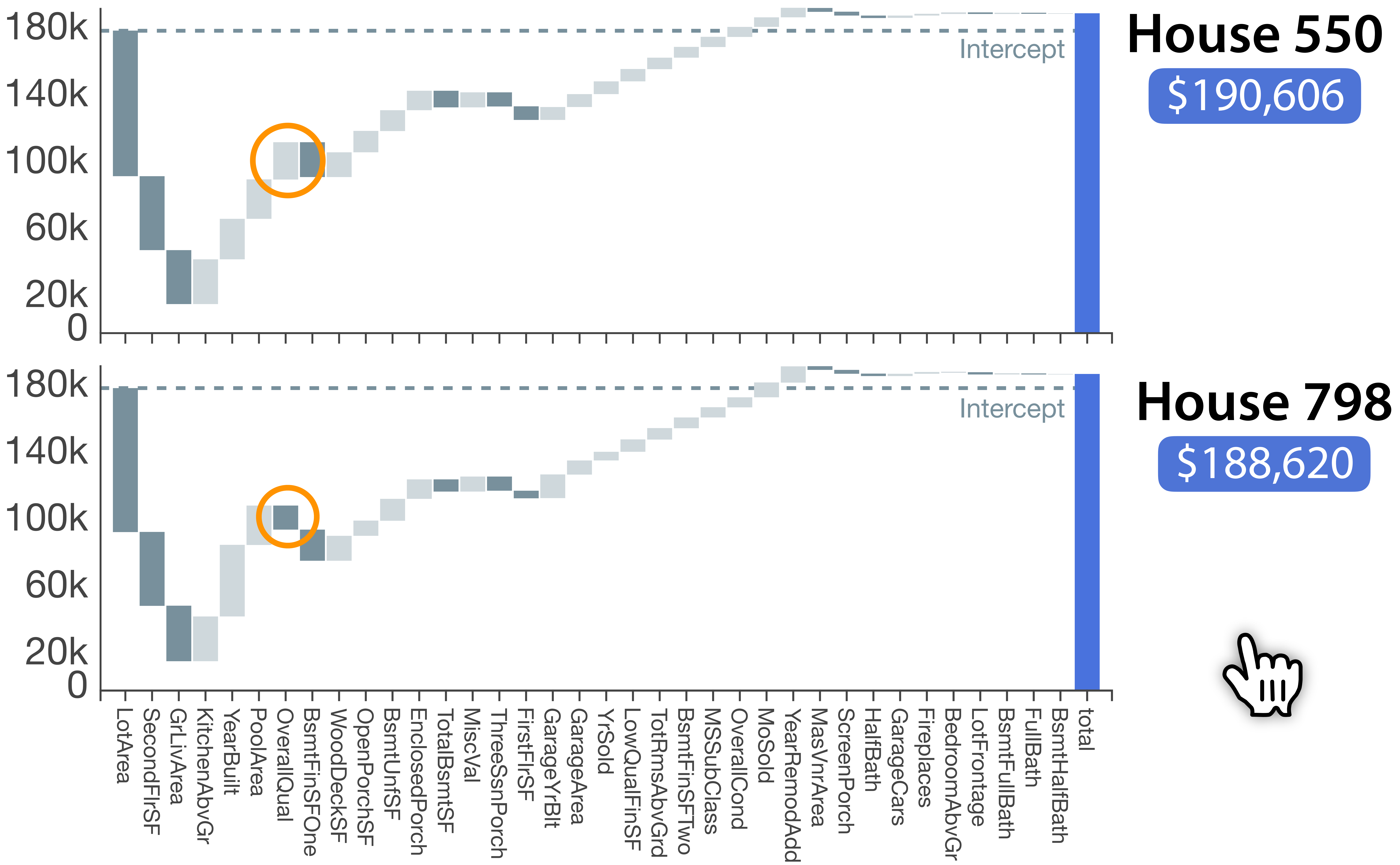

180k
140k
100k
60k
20k
0
Intercept
House 550
$190,606
House 798
$188,620
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total

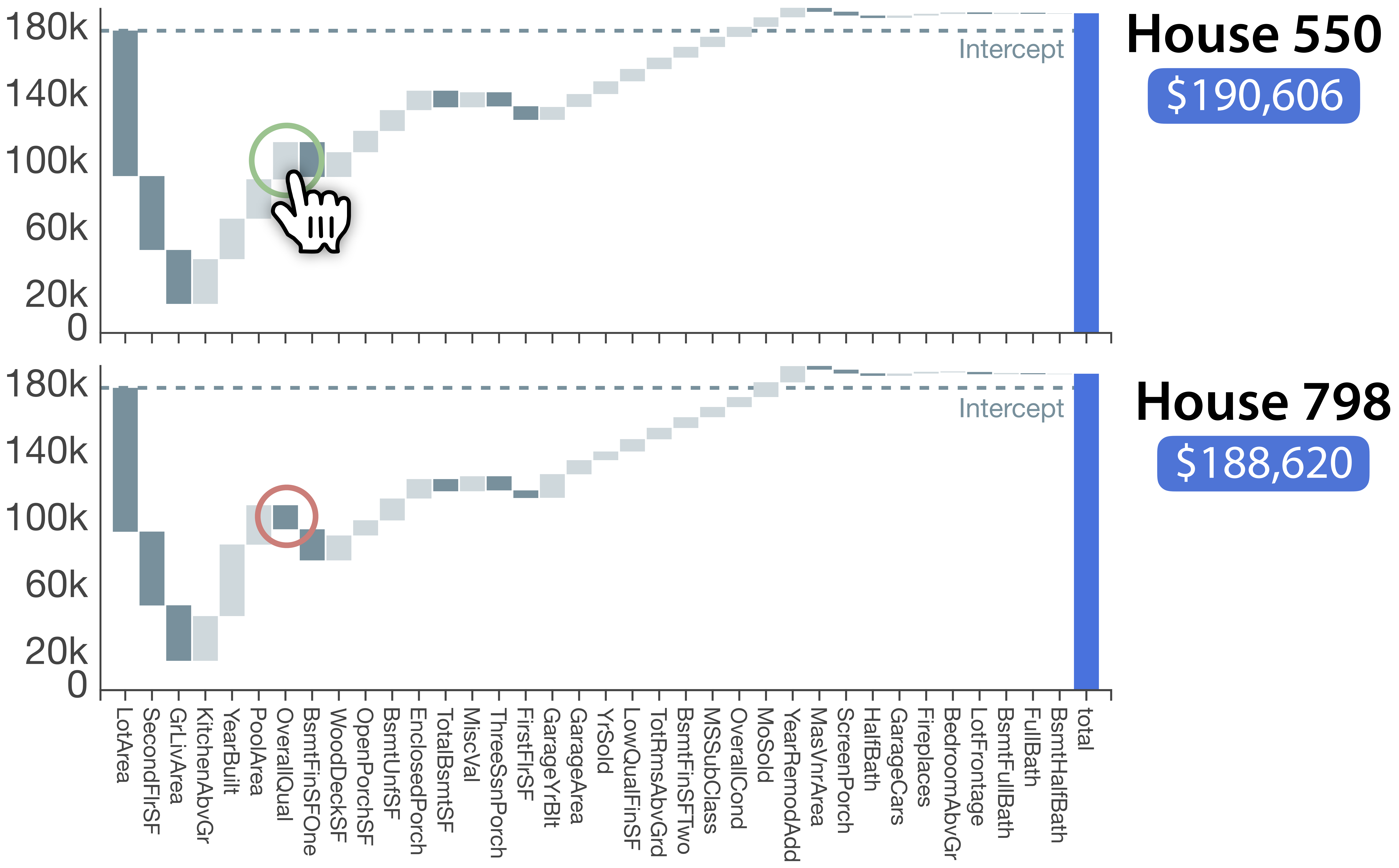


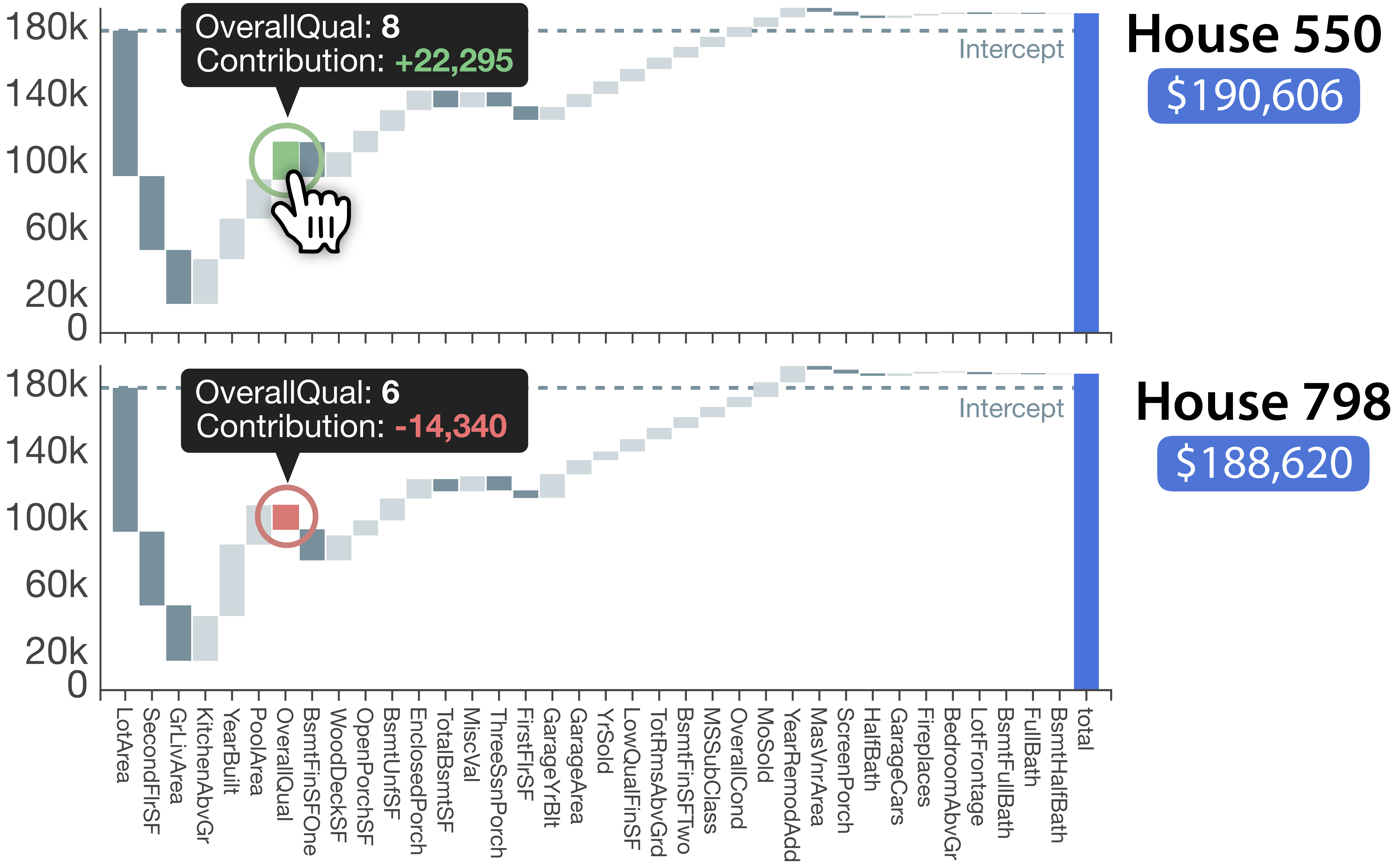
180k
140k
100k
60k
20k
0
Intercept
House 550
$190,606
House 798
$188,620
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total


180k
140k
100k
60k
20k
0
Intercept
House 550
$190,606
180k
140k
100k
60k
20k
0
Intercept
House 798
$188,620
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total


House 550
$190,606
House 798
$188,620
Intercept
Intercept
180k
140k
100k
60k
20k
0
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total


House 550
$190,606
House 798
$188,620
Intercept
Intercept
180k
140k
100k
60k
20k
0
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total

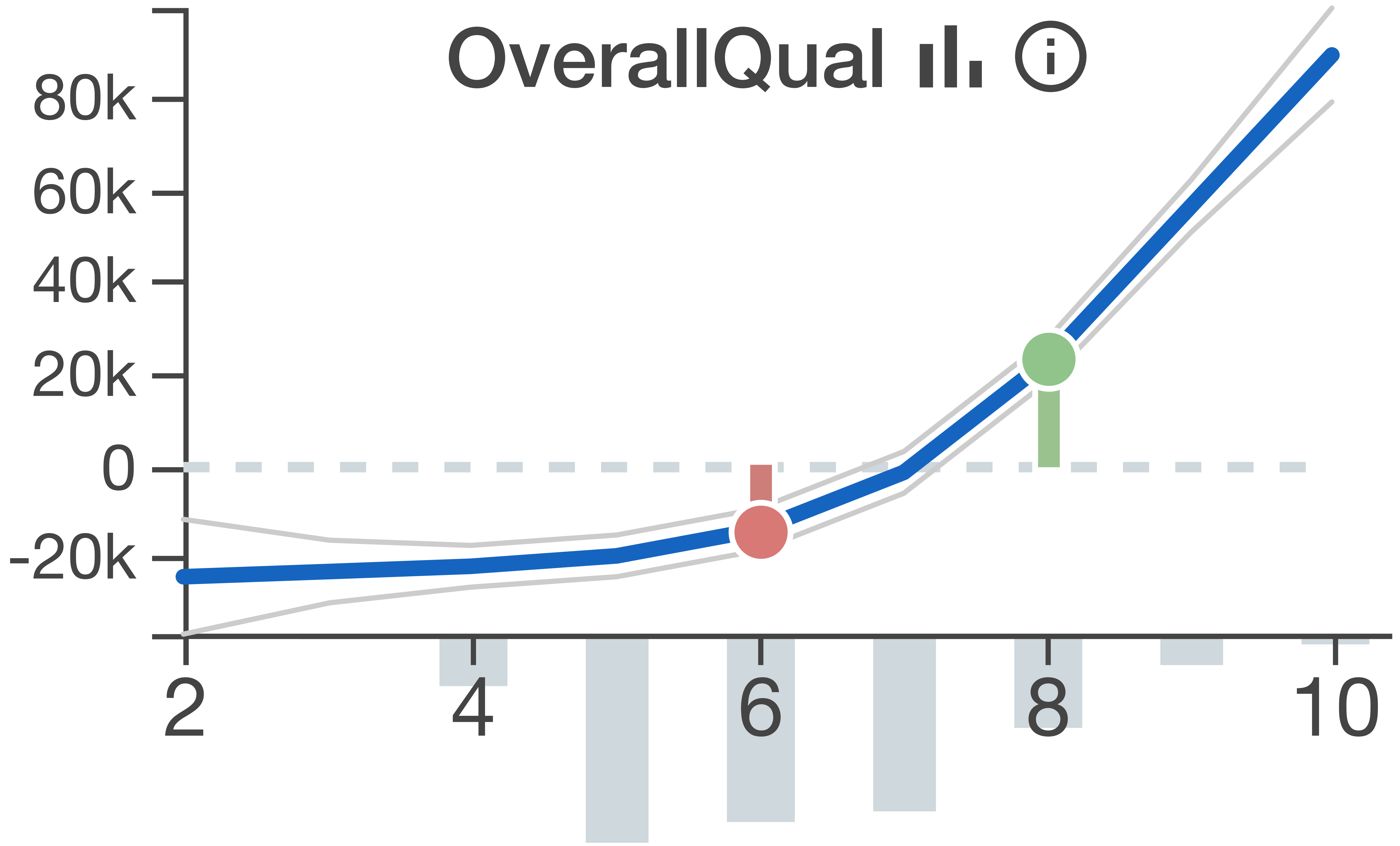

House 550
$190,606
House 798
$188,620
Intercept
180k
140k
100k
60k
20k
0
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total

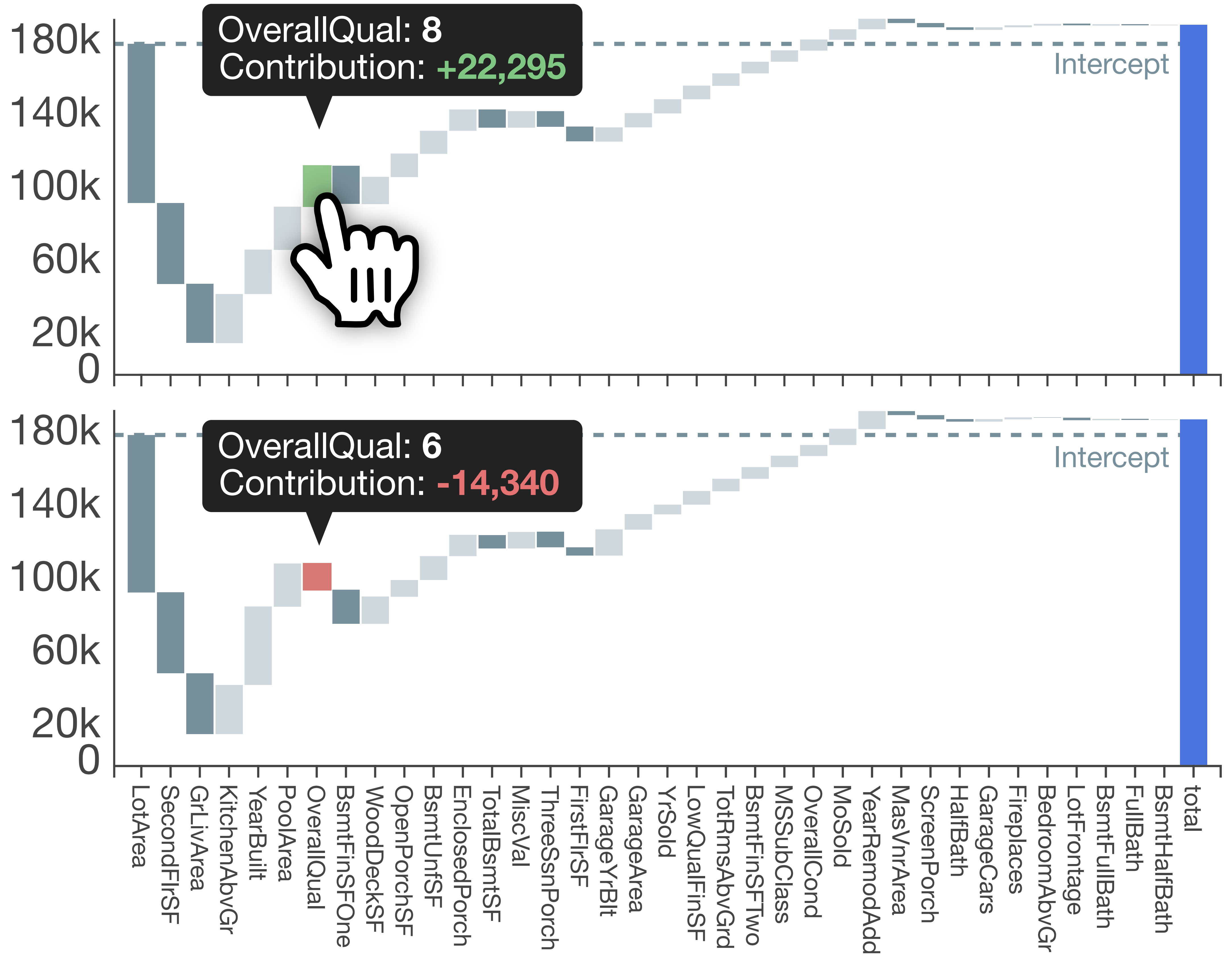


OverallQual: 8
Contribution: +22,295
Intercept
House 550
$190,606
180k
140k
100k
60k
20k
0
OverallQual: 6
Contribution: -14,340
Intercept
House 798
$188,620
LotArea
SecondFlrSF
GrLivArea
KitchenAbvGr
YearBuilt
PoolArea
OverallQual
BsmtFinSFOne
WoodDeckSF
OpenPorchSF
BsmtUnfSF
EnclosedPorch
TotalBsmtSF
MiscVal
ThreeSsnPorch
FirstFlrSF
GarageYrBlt
GarageArea
YrSold
LowQualFinSF
TotRmsAbvGrd
BsmtFinSFTwo
MSSubClass
OverallCond
MoSold
YearRemodAdd
MasVnrArea
ScreenPorch
HalfBath
GarageCars
Fireplaces
BedroomAbvGr
LotFrontage
BsmtFullBath
FullBath
BsmtHalfBath
total

House 550

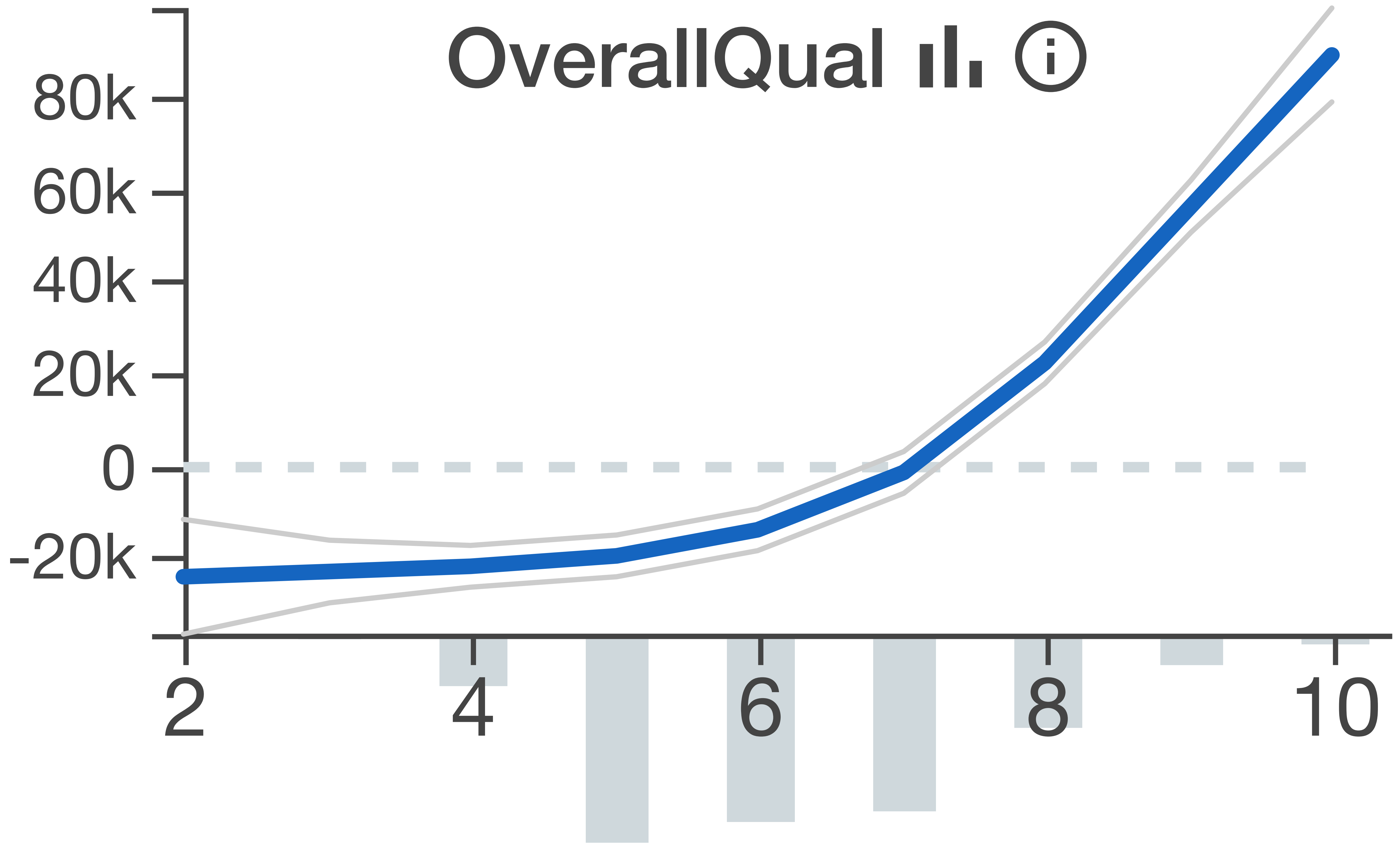
$190,606





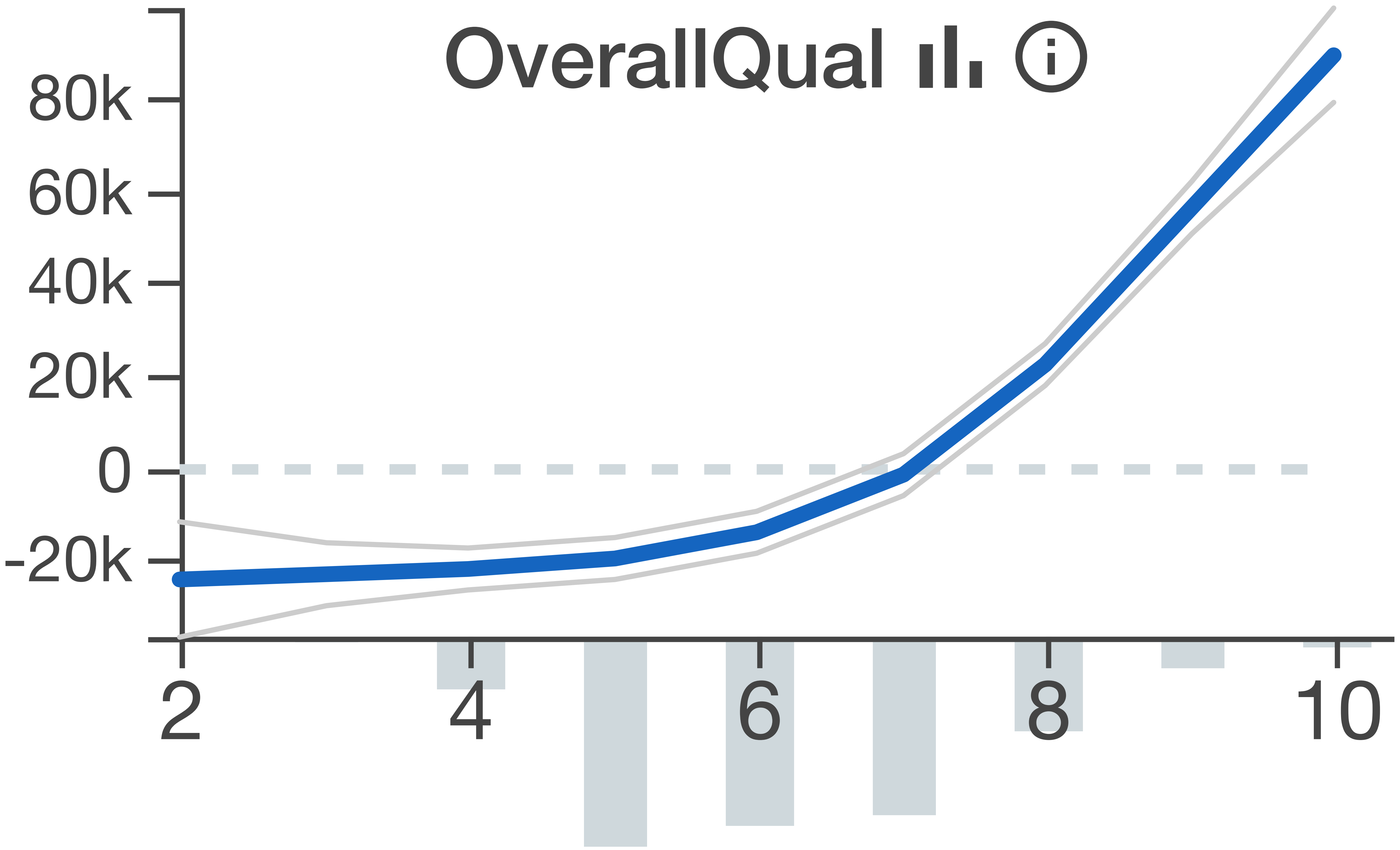
House 798

$188,620



Generalized Additive Model (**GAM**)

- ☑ Global explanation
- ☑ Easy to understand:
  - ☑ Average math skills
  - ☑ Average graphicacy
- ☑ High accuracy, realistic



Generalized Additive Model (**GAM**)

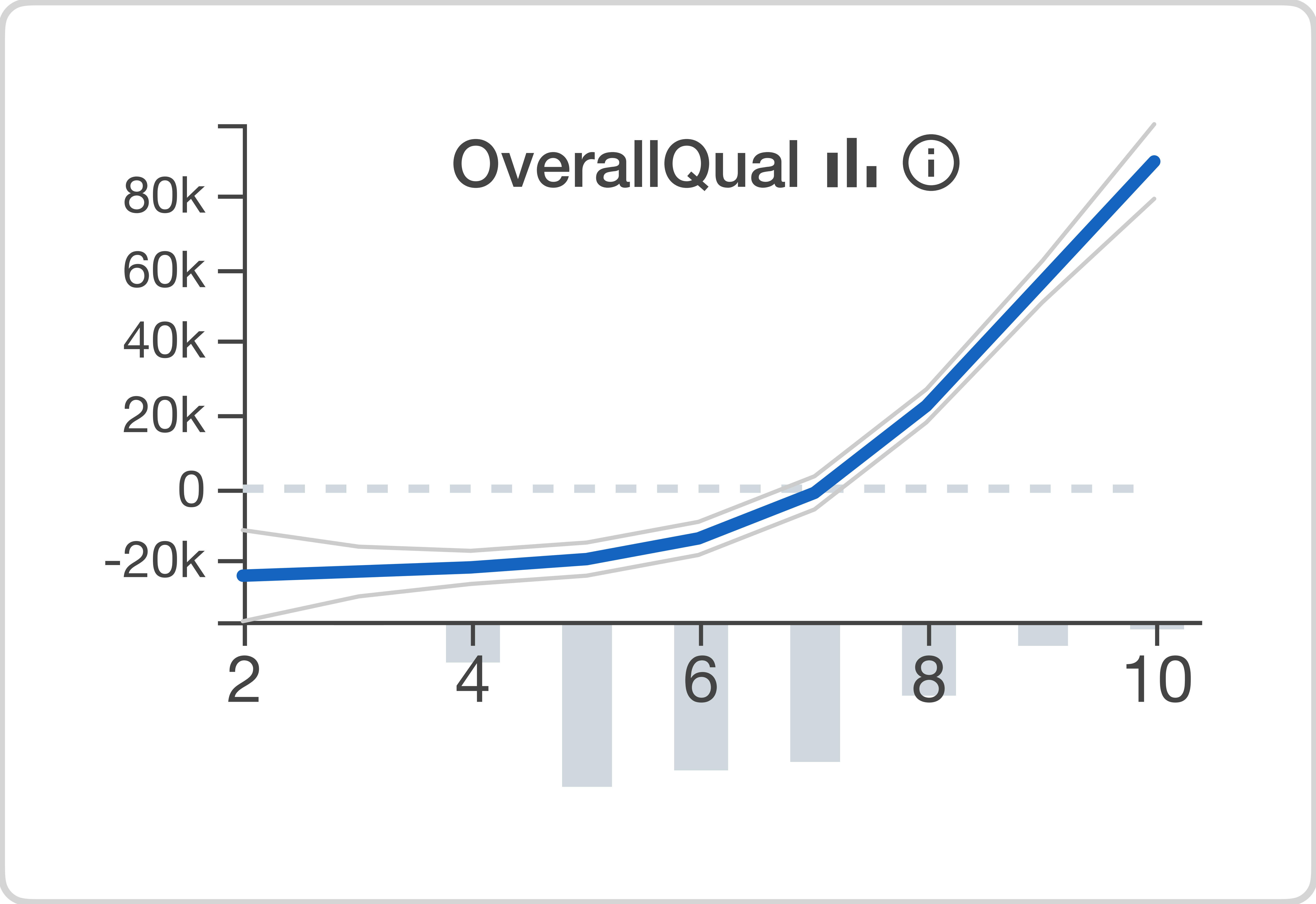
GAMs are a generalization of linear models. To illustrate the difference, consider a dataset $D = \{(\mathbf{x}_i, y_i)\}^N$ of N data points, where $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{iM})$ is a feature vector with M features, and y_i is the target, i.e., the response, variable. Let x_j denote the jth variable in feature space. A typical linear regression model can then be expressed mathematically as:

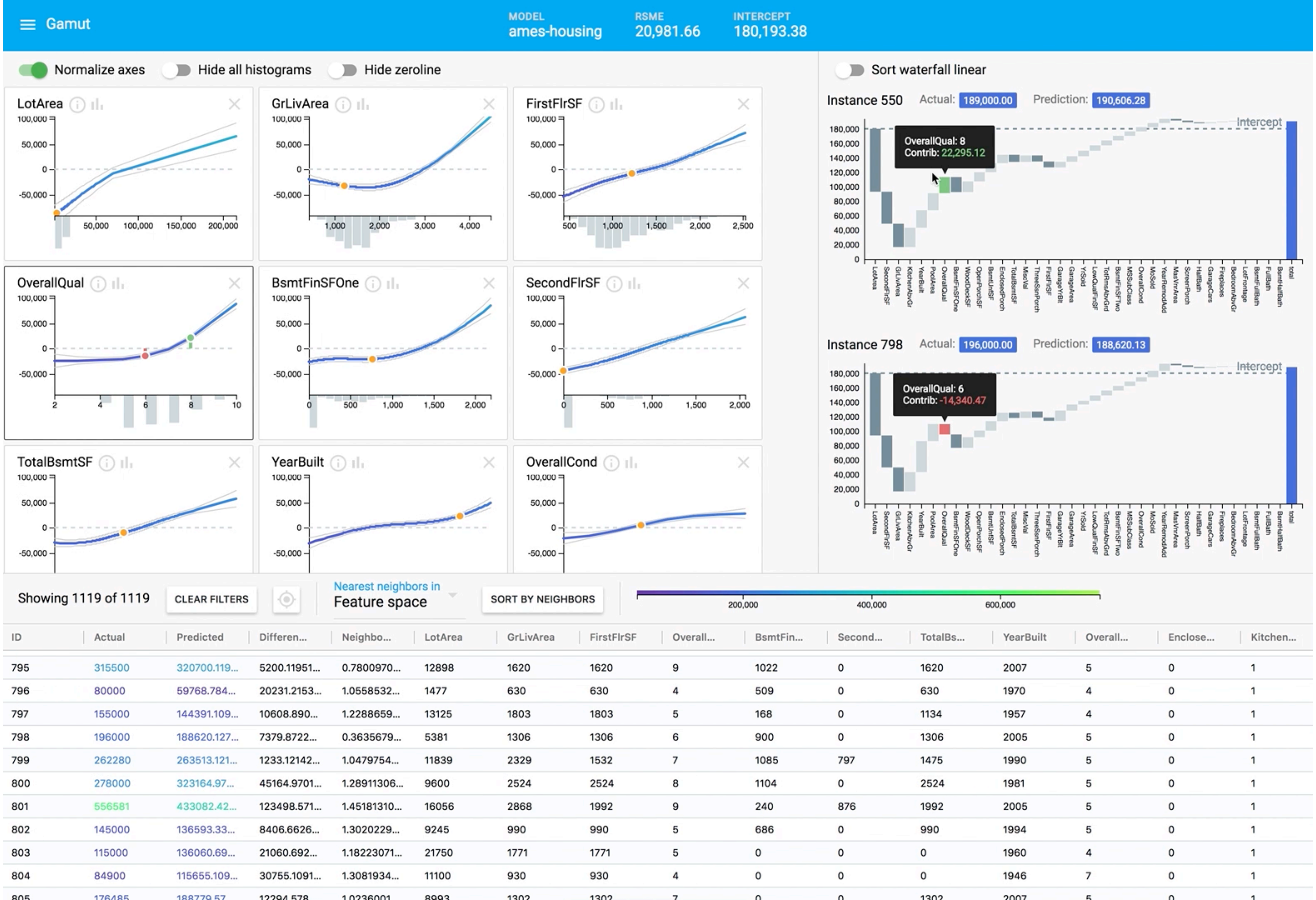
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_N x_N$$

This model assumes that the relationships between the target variable y_i and features x_j are *linear* and can be captured in slope terms $\beta_1, \beta_2, \ldots, \beta_N$. If we instead assume that the relationship between the target variable and features is *smooth*, we can write the equation for a GAM [24]:

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_N(x_N)$$

Notice here that the previous slope terms $\beta_1, \beta_2, \ldots, \beta_N$ have been replaced by smooth, shape functions f_j. In both models β_0 is the model intercept, and the relationship between the target variable and the features is still additive; however, each feature now is described by one shape function f_j that can be nonlinear and complex (e.g., concave, convex, or “bendy”) [28].


OverallQual
80k
60k
40k
20k
0
-20k
2
4
6
8
10



ID	Actual	Predicted	Differen...	Neighbo...	LotArea	GrLivArea	FirstFlrSF	Overall...	BsmtFin...	Second...	TotalBs...	YearBuilt	Overall...	Enclose...	Kitchen...
795	315500	320700.119...	5200.11951...	0.7800970...	12898	1620	1620	9	1022	0	1620	2007	5	0	1
796	80000	59768.784...	20231.2153...	1.0558532...	1477	630	630	4	509	0	630	1970	4	0	1
797	155000	144391.109...	10608.890...	1.2288659...	13125	1803	1803	5	168	0	1134	1957	4	0	1
798	196000	188620.127...	7379.8722...	0.3635679...	5381	1306	1306	6	900	0	1306	2005	5	0	1
799	262280	263513.121...	1233.12142...	1.0479754...	11839	2329	1532	7	1085	797	1475	1990	5	0	1
800	278000	323164.97...	45164.9701...	1.28911306...	9600	2524	2524	8	1104	0	2524	1981	5	0	1
801	556581	433082.42...	123498.571...	1.45181310...	16056	2868	1992	9	240	876	1992	2005	5	0	1
802	145000	136593.33...	8406.6626...	1.3020229...	9245	990	990	5	686	0	990	1994	5	0	1
803	115000	136060.69...	21060.692...	1.18223071...	21750	1771	1771	5	0	0	0	1960	4	0	1
804	84900	115655.109...	30755.1091...	1.3081934...	11100	930	930	4	0	0	0	1946	7	0	1

Contribution 3: Evaluation and Investigation

User Study

Contribution 3: Evaluation and Investigation

User Study

12 data scientists, ~1.5 hours each

User Study

12 data scientists, ~1.5 hours each

Think-aloud + answering questions:

1. data & model questions they wrote before seeing Gamut
2. prepared questions by us

Contribution 3: Evaluation and Investigation

User Study

12 data scientists, ~1.5 hours each

Think-aloud + answering questions:

1. data & model questions they wrote before seeing Gamut
2. prepared questions by us

Tutorial → Study → Interview

What we want to investigate using Gamut

Research Questions

What we want to investigate using Gamut

Research Questions

RQ1. Reasons for Model Interpretability

Why do data scientists need interpretability and how do they use it in Gamut?

What we want to investigate using Gamut

Research Questions

RQ1. Reasons for Model Interpretability

Why do data scientists need interpretability and how do they use it in Gamut?

RQ2. Global v. Local Explanations

How do data scientists use different explanation paradigms?

What we want to investigate using Gamut

Research Questions

🔬 **RQ1.** Reasons for Model Interpretability

Why do data scientists need interpretability and how do they use it in Gamut?

🌎 **RQ2.** Global v. Local Explanations

How do data scientists use different explanation paradigms?

⚡ **RQ3.** Interactive Explanations

How does interactivity play a role in explainable machine learning interfaces?

Contribution 3: Evaluation and Investigation

RQ1. Interpretability Needs and Usage 🔬

Communication is a spectrum.

"…figure out what you want emphasize and what you want to minimize. Know your audience and purpose."

Contribution 3: Evaluation and Investigation

RQ1. Interpretability Needs and Usage 🔬

Model building and debugging to boost accuracy.

"I want to understand bit by bit how the dataset features work with each other, influence each other."

Contribution 3: Evaluation and Investigation

RQ1. Interpretability Needs and Usage 🔬

Data understanding > model deployment.

“This would help me get to valuable nuggets of information, which is what [my stakeholders] are ultimately interested in.”

Contribution 3: Evaluation and Investigation

RQ1. Interpretability Needs and Usage

Hypothesis generation to help build trust.

But… eager to rationalize explanations; troublesome without healthy skepticism.

Contribution 3: Evaluation and Investigation

RQ2. Global v. Local Explanations 🌎

Contribution 3: Evaluation and Investigation

RQ2. Global v. Local Explanations 🌎

Contribution 3: Evaluation and Investigation

RQ2. Global v. Local Explanations 🌎

Global

features + model

Contribution 3: Evaluation and Investigation

RQ2. Global v. Local Explanations 🌎

Global
features + model

Local
single instances

Contribution 3: Evaluation and Investigation

RQ2. Global v. Local Explanations 🌎

Global
features + model

Local
single instances

ML novice
[1-3 years]

Contribution 3: Evaluation and Investigation

RQ2. Global v. Local Explanations

Global
features + model

ML familiars
[3-5 years]

Local
single instances

ML novice
[1-3 years]

Contribution 3: Evaluation and Investigation

RQ2. Global v. Local Explanations 🌎

ML experts
[5+ years]

Global
features + model

ML familiars
[3-5 years]

Local
single instances

ML novice
[1-3 years]

Contribution 3: Evaluation and Investigation

RQ3. Interactive Explanations ⚡

RQ3. Interactive Explanations ⚡

Primary mechanism for exploring, comparing, and explaining predictions

RQ3. Interactive Explanations ⚡️

Primary mechanism for exploring, comparing, and explaining predictions

Converse with a model

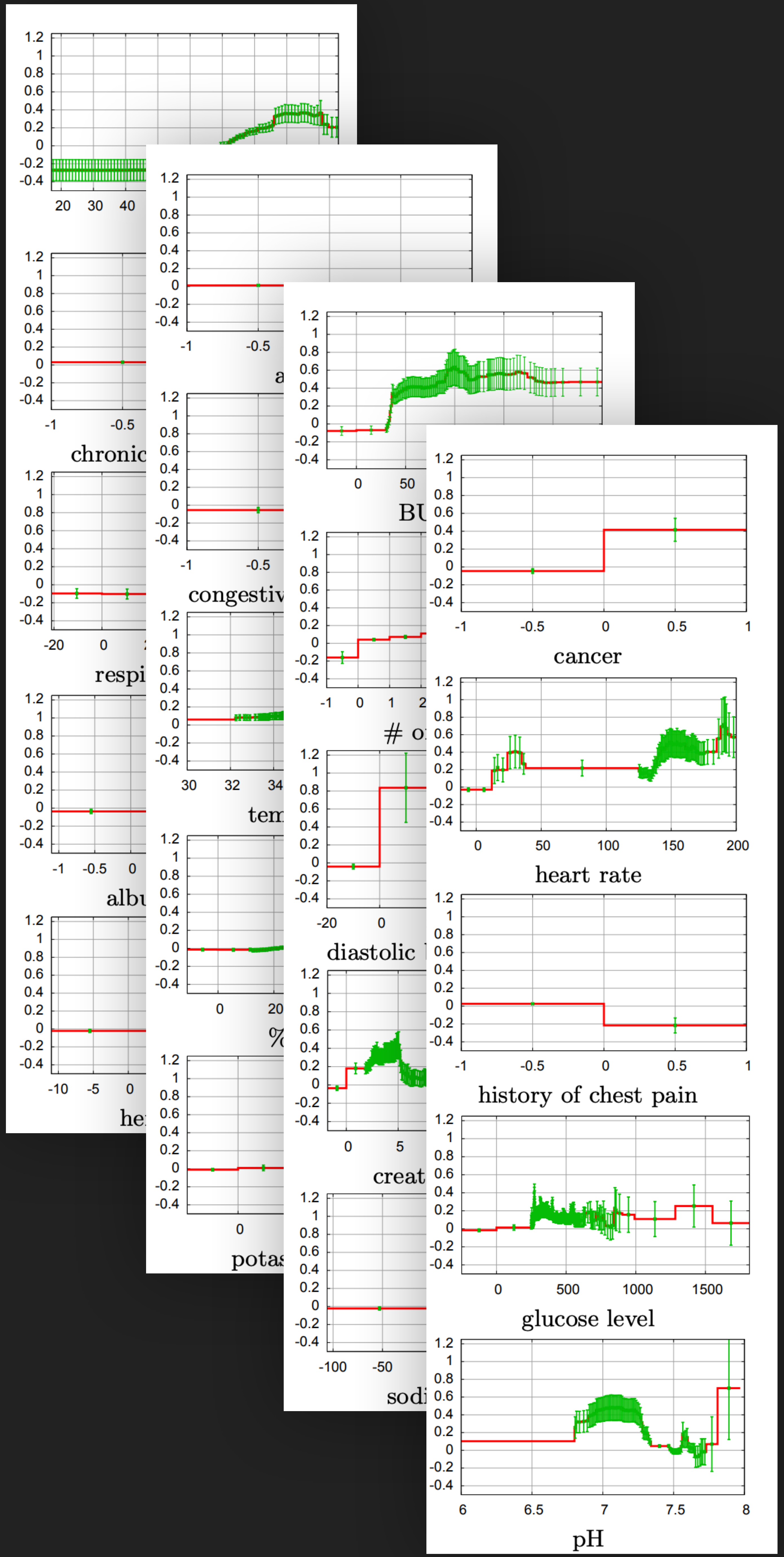
Contribution 3: Evaluation and Investigation

RQ3. Interactive Explanations ⚡

Primary mechanism for exploring, comparing, and explaining predictions

Converse with a model

Could not conceive of non-interactive



Takeaways

Takeaways

🔬 **Consider interpretability capabilities for your interfaces**
Interpretability is ***not a singular, rigid concept***

Takeaways

🔬 **Consider interpretability capabilities for your interfaces**
Interpretability is ***not a singular, rigid concept***

🌎 **Tailor explanations for specific audiences**
Balance ***simplicity*** and ***completeness***

Takeaways

🔬 **Consider interpretability capabilities for your interfaces**

Interpretability is ***not a singular, rigid concept***

🌎 **Tailor explanations for specific audiences**

Balance ***simplicity*** and ***completeness***

⚡ **Design and integrate effective interaction**

Interaction key to ***realizing interpretability*** & solidify model understanding [Weld & Bansal, 2018]

Gamut
A Design Probe to Understand How Data Scientists Understand Machine Learning Models

bit.ly/gamut-chi

📝 paper 🎥 video 📰 blog 📺 slides

Thanks!


Fred Hohman
@fredhohman
Georgia Tech


Andrew Head
UC Berkeley


Rich Caruana
Microsoft Research


Rob DeLine
Microsoft Research


Steven Drucker
Microsoft Research

Georgia Tech · NASA · Berkeley · Microsoft Research

extra slides

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

target

General **Linear** Model

Features

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

intercept

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

Slope terms

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

$$\Downarrow$$

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

$$\Downarrow$$

Generalized **Additive** Model

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$$

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

$$\Downarrow$$

Generalized **Additive** Model

Shape Functions

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$$

General **Linear** Model

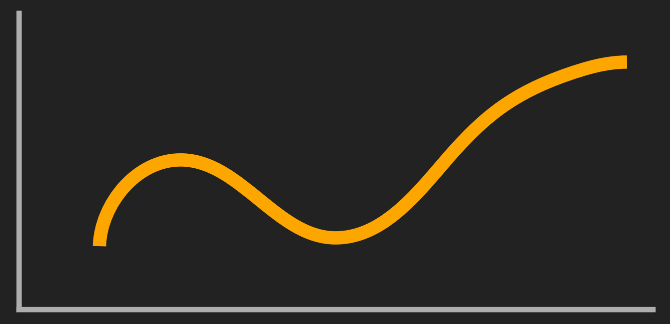
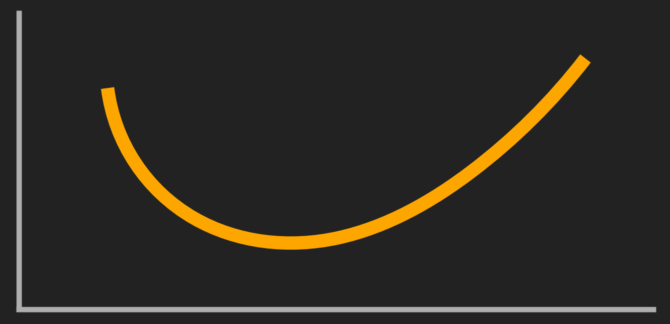
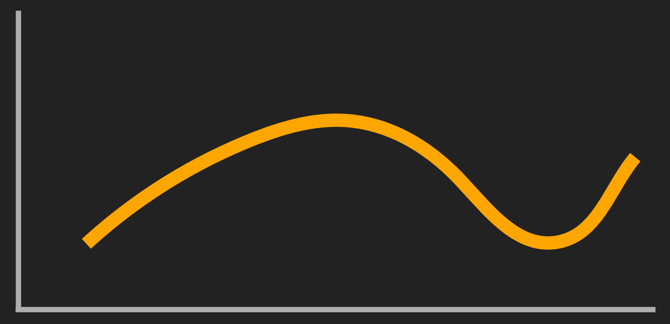
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

$$\Downarrow$$

Generalized **Additive** Model

Shape Functions

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$$

General **Linear** Model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

$$\Downarrow$$

Generalized **Additive** Model

Shape Functions

nonlinear, or "bendy"
[Jones & Almond, 1992]

$$y = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_n(x_n)$$