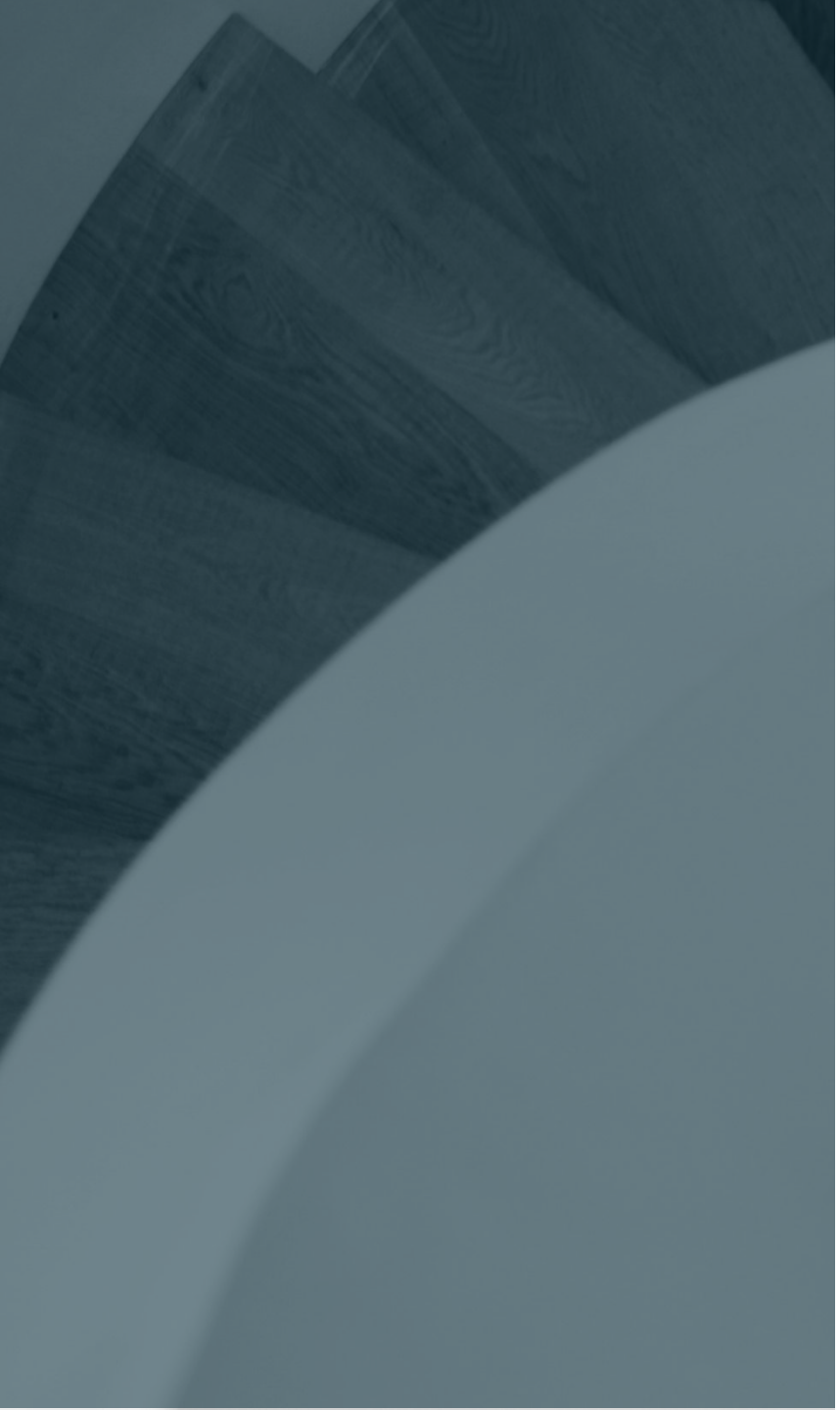
Proposal Success Indicators

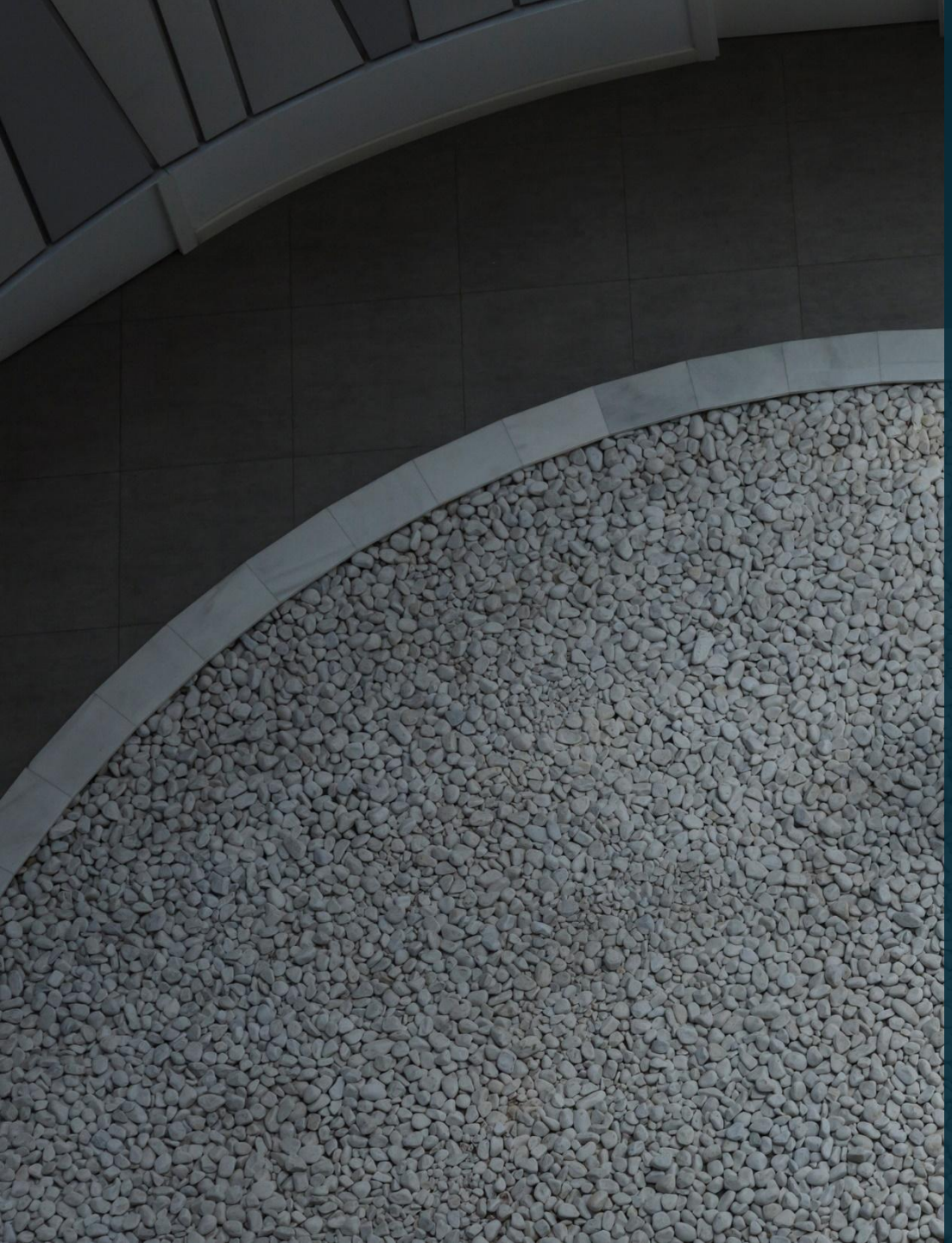
Data Alignment & Sign-off

December 2024
By: Nicole Quevillon



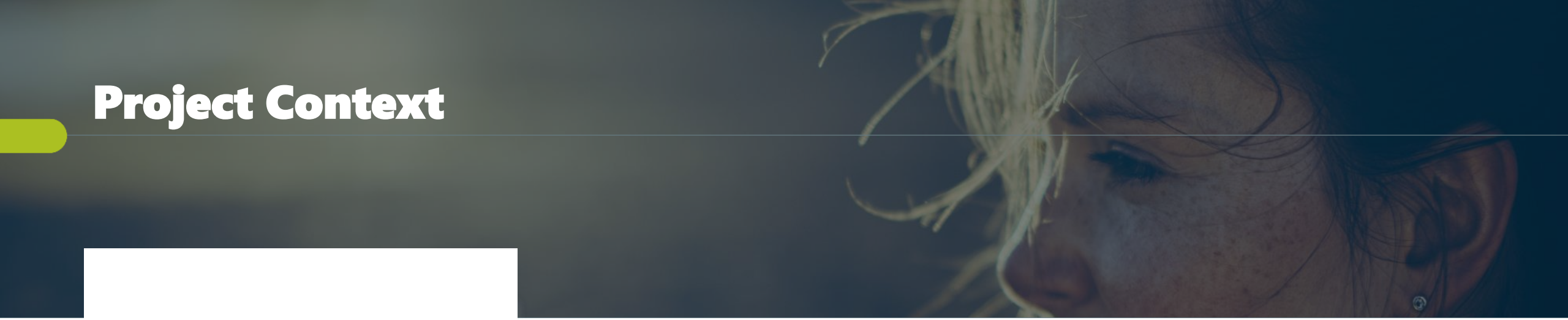
contents

This presentation summarizes the adjustments made to the proposal data. The purpose of this deck is to gain alignment on these adjustments with the lead project stakeholders before moving forward with the analysis. As such, this is a detailed technical document.



Introduction

Project Context



The Assessment & Restoration team (A&R) is looking to identify areas where they can focus their proposal efforts to maximize revenue impact.

Over the past three years, the team has submitted over 700 proposals per year with a success rate slightly below 50%.

With a goal of growing their 2025 pipeline revenue by 12%, the A&R team needs to decide how they will shift their proposal efforts during the year to hit this target.

Rather than indiscriminately increasing proposal submissions, the team can use proposal resources more effectively by prioritizing proposal types focused on.

Project Purpose

The purpose of this project is to identify key areas where proposal success rates are consistently higher or lower than the average.

Machine learning will be used to find trends among the thousands of proposal submitted by A&R over the past five years.

This analysis will identify:

1. **Proposal types with a higher success rate where the team can focus efforts for increased revenue impact.**
2. **Proposal types with a lower success rate where current offerings could be made more competitive.**

Data Adjustments for All Proposals

Outcome

This is the key column we will be using to identify proposal success

We will be creating a binary classification model, so only Wins and Losses will be included in the model.

Adjustment: Any proposal submitted more than one year before the date that the data was downloaded (October 1, 2024) that is still listed as 'Unknown' was assumed to be a loss. All other unknowns were removed from the data.

Region

The first two digits of the Client column were used to create a Region indicator. These were then grouped into larger regionalities

Client	Region (ver 1)	Region (ver 2)
	Created by taking the first two digits of the Client column. In condo corps, these digits indicate the Land Registry Office where the corporation was incorporated	Land Registry Office-based regions were then grouped together into the major regions of: TOR, GH, EO, HUR, Other and Unknown
OC0754	OC	EO
Finn Projects	Fi	Unknown
TS2118	TS	TOR
YR1252	YR	TOR
CAPREIT	CA	Unknown

Region

Some additional cleansing was done to the final region groupings

Region	# of Instances	Region Long Name	Includes municipalities such as:
TOR	1205	Toronto	Toronto, Pickering, Ajax
EO	331	Eastern Ontario	Ottawa, Kingston, Coburg
YR	262	York Region	Aurora, Newmarket
HUR	243	Huronia	Barrie, Owen Sound, Bracebridge
PC	174	Peel	Mississauga, Brampton
GH	91	Golden Horseshoe	Milton, Hamilton, St. Catherines
LON	13	London	London, Woodstock, Stratford
GR	12	Grand River	Kitchener, Guelph, Brantford
WIN	6	Windsor	Windsor, Sarnia, Chatham
NW	0	North West	Thunder Bay, Kenora, Cochrane
Unknown	1251		Catchall for anything that doesn't match

Adjustment: These were grouped into an 'Other' category due to the low # of instances in each

Project Types

Similar SOW categories were grouped to make a Project Type Indicator

Project Type Group	SOWs Included	Count of Proposals
RFS	['Update RFS' 'Class 2 RFS' 'Class 3 RFS' 'Class 1 RFS']	1,370
BCA	['BCA & RFS' 'BCA']	291
Consult	['Consult']	286
PCA	['PCA']	269
Envelope	['Stairs' 'Envelope' 'Balcony' 'Windows' 'Doors' 'Sealants' 'Structural']	204
Roofing	['Roofing']	154
Mechanical	['Mechanical']	116
Garage	['Garage']	98
PA & Class 1 RFS	['PA & Class 1 RFS']	88
Other	[PA1, C1, PA2, C2 RFS with C3 Fee, PPA, PA Year 1, PA Follow-Up, Renovation, Pavement, PA Year 7, PA Year 2, PCA/ESA, PCA & ESA, Landscaping, Accessibility Audit, Bulletin 18, Energy Audits, Electrical]	179

This classification is what was used to categorize projects by either RFS or ALL

Scope of Work

Within SOW, anything with under 20 instances was grouped into an 'Other' or 'Renovation' category

'Other' category includes:

- Class 3 RFS 16
- PA Year 7 6
- C2 RFS with C3 Fee 5
- PA Follow-Up 4
- BCA & RFS 4
- PPA 4
- Electrical 1
- PA Year 1 1
- PA Year 2 1
- Accessibility Audit 1

'Renovation' category includes:

- Windows 13
- Renovation 12
- Landscaping 12
- Sealants 6
- Stairs 3
- Doors 1

This is the only adjustment applied to the All proposal dataset that was not also applied to the RFS dataset. All sow types were kept for RFS

RFPs

There were only 42 instances of proposals linked to RFPs and they were quite different from the average proposal, so these were removed from the data

	Proposal Outcome		Proposal Fee	
	Non-RFP	RFP	Non-RFP	RFP
Count	3,056	**96**	3,056	96
Mean	51%	**36%**	$11,371	**$148,450**
Median			$3,600	$44,169

Data Adjustments for Reserve Fund Study Proposals

RFS Data

Due to the process used by the team to record building details for Reserve Fund Study (RFS) proposals, there is significantly more data available for this proposal type.

The data outlined in this section is only available for RFS proposals. Due to this, there will be two sets of data and models, one for RFS, and one for all other proposals.

The data outlined in the previous section is also applicable to RFS proposals.

Unit Counts

The NO. UNITS column was messy, so the unit count was extracted differently for the various building types

Commercial

Takes the numeric digits before an instance of com(?:mercial) or ind(?:ustrial) in the NO. UNITS column

POTL

Takes the numeric digits before an instance of 'POTL' in the NO. UNITS column

Residential

Any plain numeric value in the NO. UNITs column was assumed to be the number of residential units.

Any numeric values before the word res(?:idential) in the NO. UNITS column were also used.

Building Type Flags Created

Flags were created to indicate what type of building the proposal was for

Building Type	Classification Rules	Count of Proposals
High-Rise	Any proposal for a building with more than 3 storeys. If there was no storey data, anything with more than 68 units (this was the 25th percentile for high-rises we had been able to identify already) was assumed to be a high-rise	422
Townhouse	Any proposal with a value in the NO. TOWNHOUSE BLOCKS column (with the exception of the value '-') is flagged as a townhouse.	388
Other Residential	Any property not classified with the existing rules, but which has residential units (as defined on previous slide)	144
Commercial	Any proposal where the string 'com' or 'ind' appears in any of the following columns NO. UNITS, NO. STOREYS or NO. TOWNHOUSE BLOCKS is flagged as commercial	140
POTL	Any proposal with a non-null POTL unit count (as defined on previous slide)	110
Mixed Use	Any proposal flagged as commercial, but which also has residential units (as defined on previous slide)	29
Unknown	Catchall for any proposal we were unable to classify using the above rules	171

Number of Storeys

The NO. STOREYS column required significant cleaning

Adjustments

1. If the entry contained no numbers, it was set to null
2. If the entry is only numbers, leave as is
3. If there is a mix of letter and numbers in the entry, grab only the numeric digits that appear before the word 'storey'
4. If the building is POTL or Townhouse, assume it is 2 storeys

Nulls: 538

Not nulls: 832

Despite these adjustments, there are still quite a few nulls in this column. This will not be useable in modelling.

Example of values that were not able to be converted:

```
['-' '2 & 4' '2 or 3' '1+basement' 'Tower 1 - 10\nTower 2 - 14' '2 and 1'
 'Multi' '4  and 3' '21 and 5' ' 2 / 3' '7 and 8' ' 1 - 2 ' '1 / 2'
 '.2 / 3' '6 x 3' ' 25 & 12' '2 x 4' '2 x 3' '2 x 23' 'Various Stories'
 '4 / 7' '9 + TH' '13/9' '15 / 32' '14/49/14/39' '6\n6' '34/25' '2 x 12'
 '1 and 2' '7 x 2' '1 and 3' '2/3' '33/13/5' '12 + 8' '3/1' '11/29'
 '2 x 6' '19 + 12' '2 x12' '1 to 2' '32\n41' '2 x4' 'Queen-18 & Sparks-6'
 'Queen-18 & Sparks-6 Parking Only' 'Queen-18' 'Queen-18 & Sparks-Hotels'
 '2 x 15' '17, 14, 12' '2 buildings x 3' '44\n45' '16 blocks'
 '10 Apartment\n2 Rec Building' '16\n195 HR units'
 '18 Davenport Rd - 16\n4455 Bathurst St - 14\n193 Wilson Ave - 5'
 '2 towers x24' '65 Bldgs' '6; 7; 3' '10-Buildings'
 '22 (48 tower)\n20 (50 Tower)' '38 each (2 towers)' '35\n16'
 '5 (2 underground & 3 above ground)' 'Bldg A & B - 20\nBldg C to F - 16'
 '15 (condo & commercial)' 'Bldg A - 1\nBldg B & C - 2'
 '2buildings - 7 floors' '3+1 raised basement\n2+1 raised basement'
 '57 &  67' '3 Bldgs - 14' '9\n6\n9' '9\n9' '3+basement' '31\n31' '?\n33'
 'Logan Ave - 22\nMutual St - 4' '2 Bldgs - 44,000 SF & 4,000 SF'
```

Final Datasets

Data Available for All Proposals Modelling

This is what we will be trying to understand & predict

Column	Description
Region	Region that the building is located in
SOW (scope of work)	Project type being proposed
Project Type	Groupings of the SOW
Fee_probable	The revenue that would be expected if the proposal was successful
Issued_date	The date that the proposal was submitted to the prospective client
Outcome	**The outcome of the proposal**

Data Available Only for RFS Proposal Modelling

Column	Description
Prev_rfs	Whether Egis conducted the client's previous RFS
reminder	Whether the proposal is being sent as a reminder (rather than the client reaching out to ask for a proposal)
building_type	Indicates whether the building is a townhouse, highrise, commercial etc
Building_age	Age of building on the date that the proposal is submitted
units	Number of units in the building

Nicole Quevillon

nicoleaquevillon@gmail.com