



איך להתמודד עם בקרת נתוני ולא לטבוע בים של בקרות לא אפקטיביות?

אירנה ויירו
אפריל 2024



מה?

למה?

איך?



למה?!



דרישות רגולטוריות

ממשל תאגידי

אקטואר ממונה ואקטואר ראשי

1

SOX

2

חוזר דרישות תיעוד מודלים, הנחות מחקרים

3

SOLVENCY/ORSA

4

IFRS17

5



חוק

GIGO

PRECISE + PRECISE = SLIGHTLY LESS
NUMBER + NUMBER = PRECISE NUMBER

PRECISE × PRECISE = SLIGHTLY LESS
NUMBER × NUMBER = PRECISE NUMBER

PRECISE + GARBAGE = GARBAGE
NUMBER + GARBAGE = GARBAGE

PRECISE × GARBAGE = GARBAGE
NUMBER × GARBAGE = GARBAGE

$\sqrt{\text{GARBAGE}}$ = LESS BAD
GARBAGE

$$(\text{GARBAGE})^2 = \text{WORSE GARBAGE}$$

$$\frac{1}{N} \sum (N \text{ PIECES OF STATISTICALLY INDEPENDENT GARBAGE}) = \text{BETTER GARBAGE}$$

$$(\text{PRECISE NUMBER})^{\text{GARBAGE}} = \text{MUCH WORSE GARBAGE}$$

$$\text{GARBAGE} - \text{GARBAGE} = \text{MUCH WORSE GARBAGE}$$

$$\frac{\text{PRECISE NUMBER}}{\text{GARBAGE} - \text{GARBAGE}} = \text{MUCH WORSE GARBAGE, POSSIBLE DIVISION BY ZERO}$$

GARBAGE = 0 = PRECISE
NUMBER



בקרת נתונים משפרת את הבטחון בקבלת החלטות



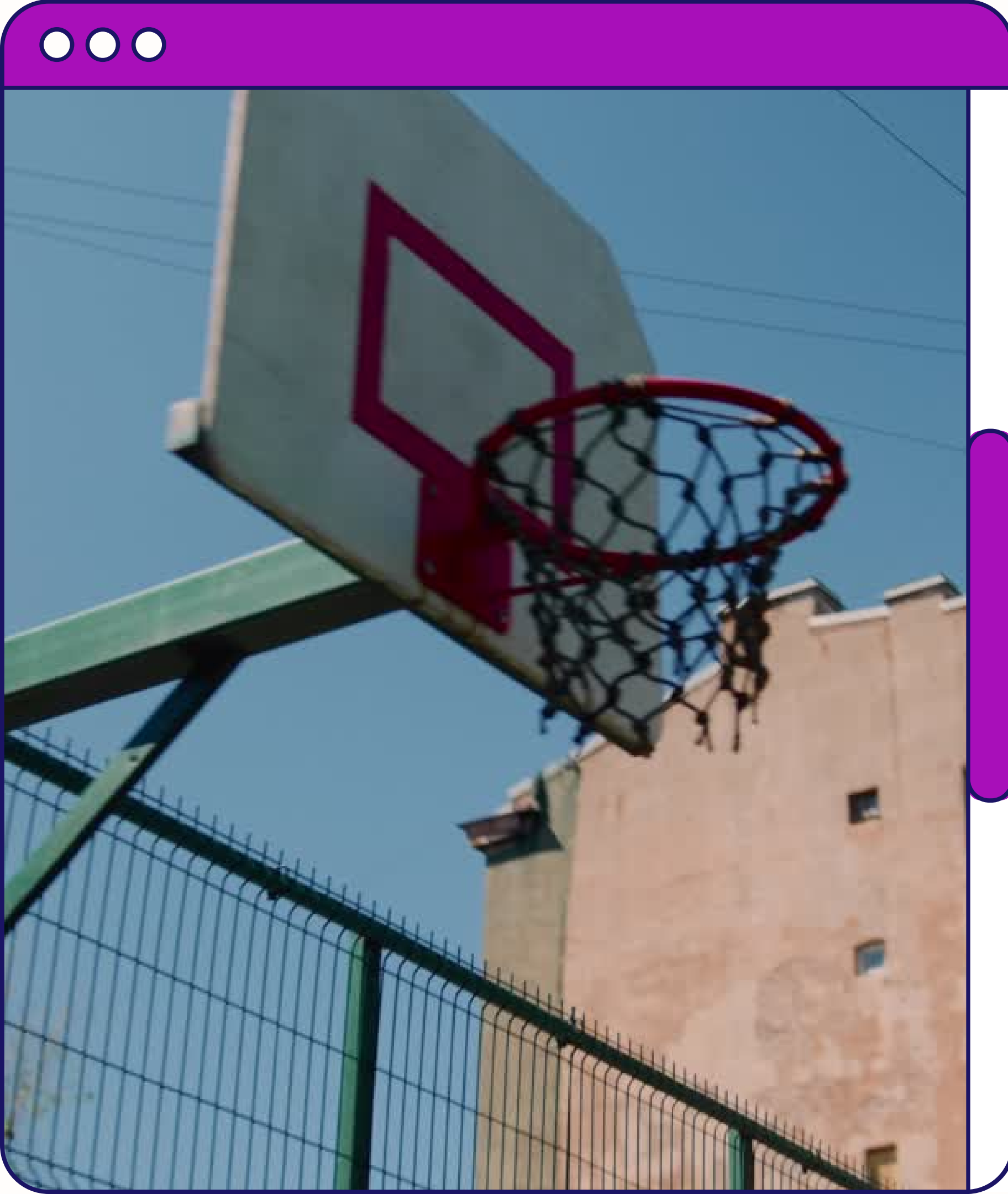
בקרת נתונים חוסכת זמן עיבוד במודל



בקרת נתונים תורמת לעמקת הבנה עסקית והכרות עם המגמות



מה?



Accuracy

נכונות

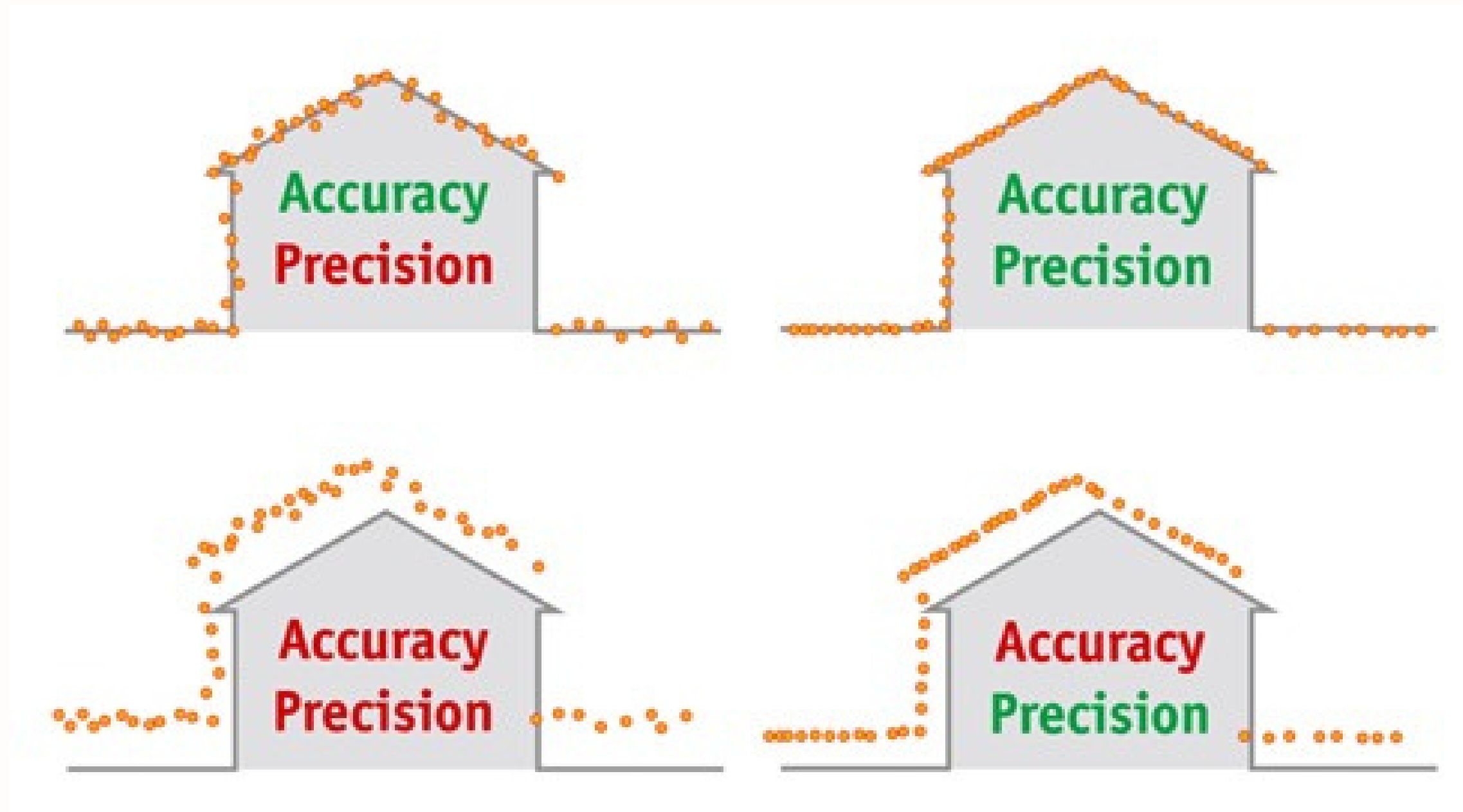


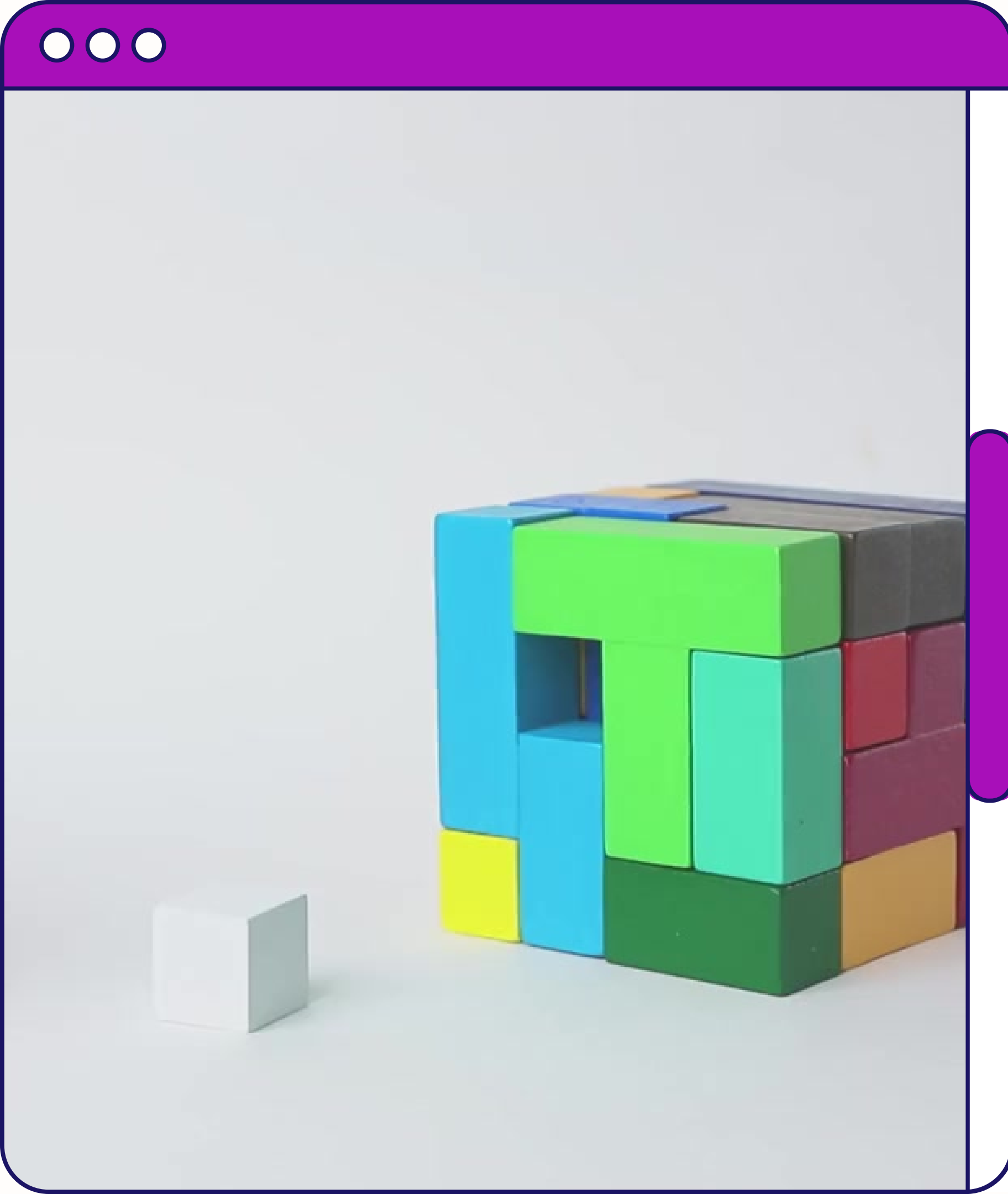
Precision

דיוק

Accuracy VS Precision

נכונות \neq דיוק





Completeness

שלמות



Timeliness

עדכניות



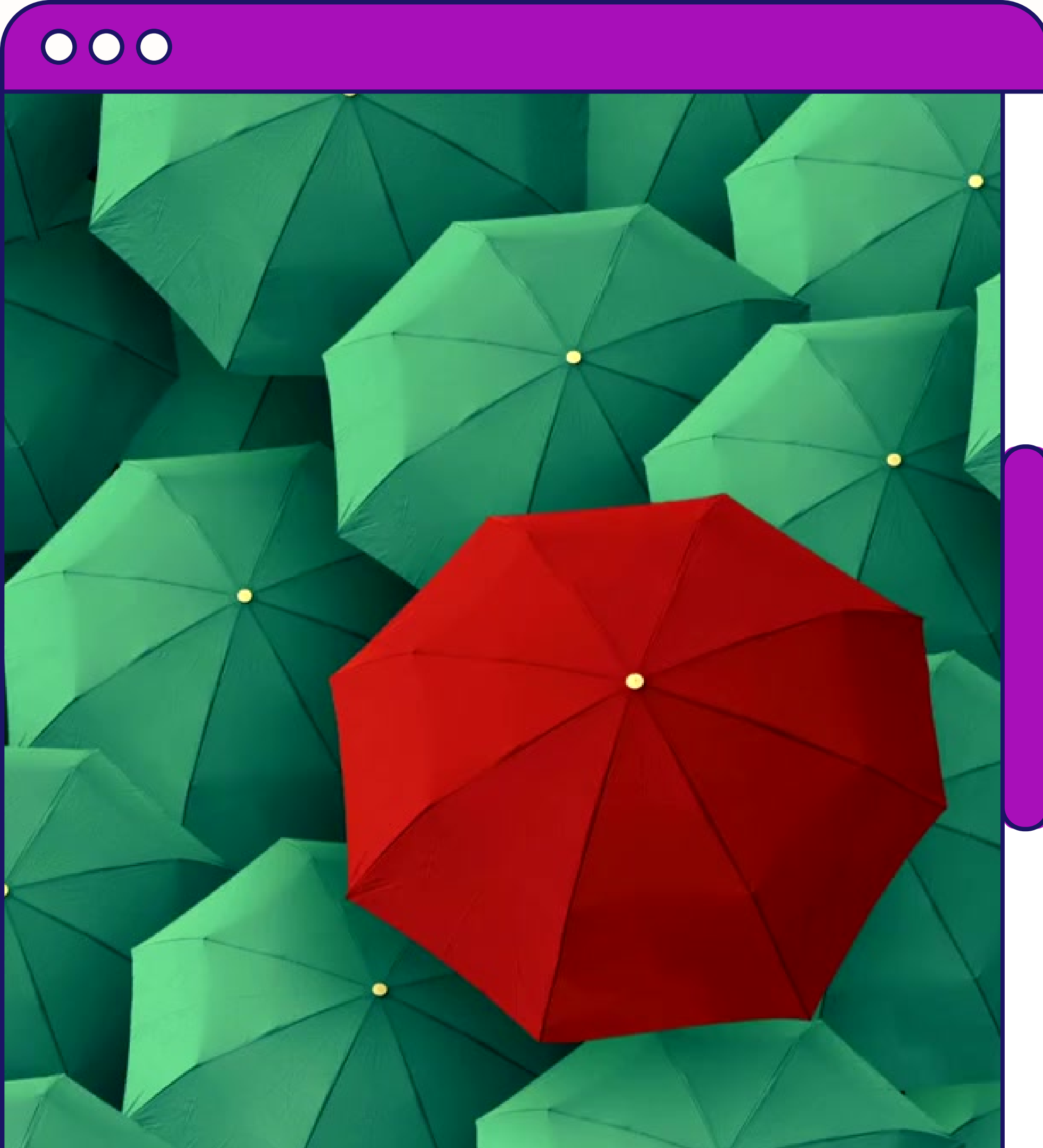
Timeliness

עדכניות

online VS batch 1

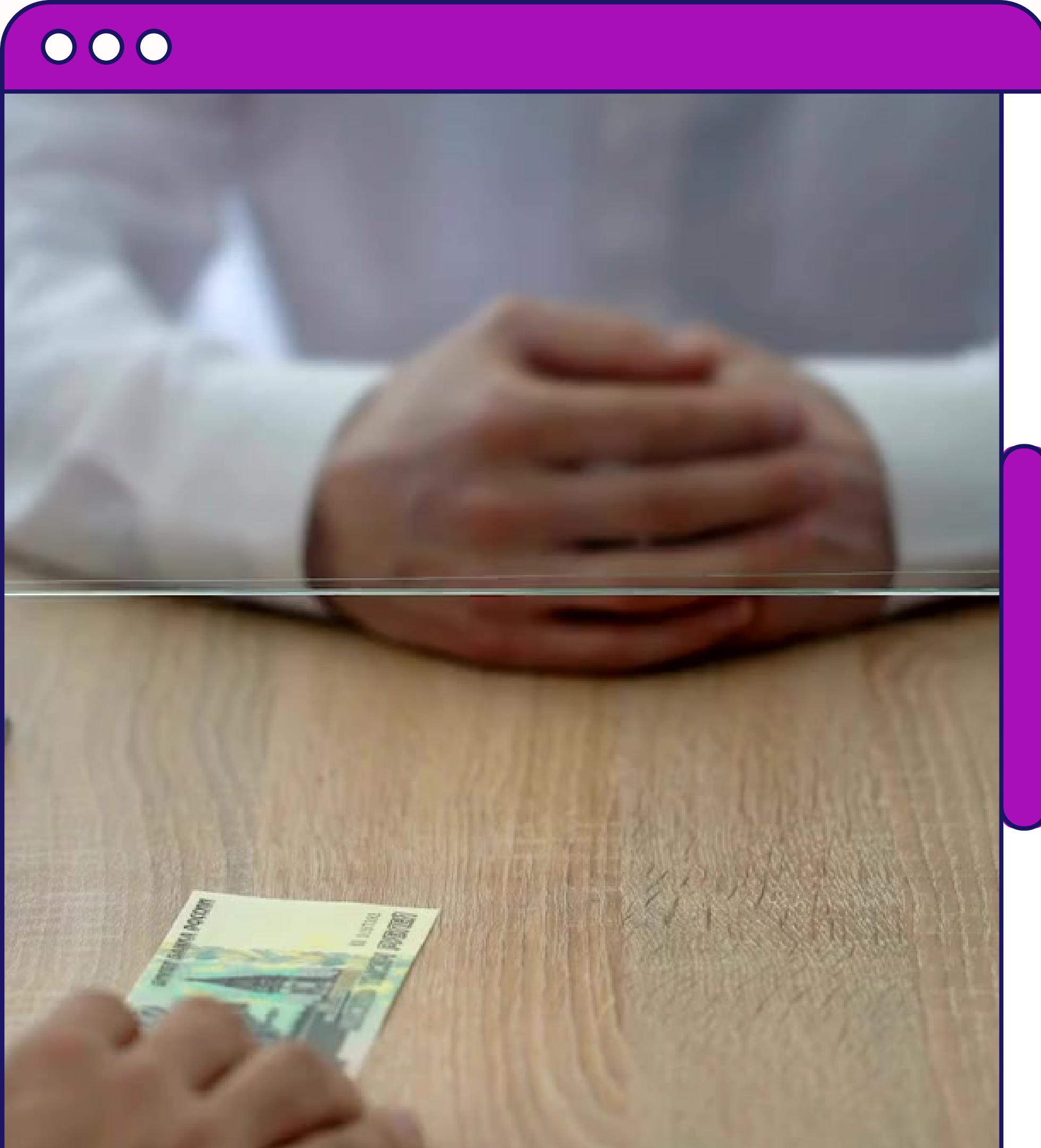
שחזור נתונים 2

שינויים רטרואקטיביים 3



Uniqueness

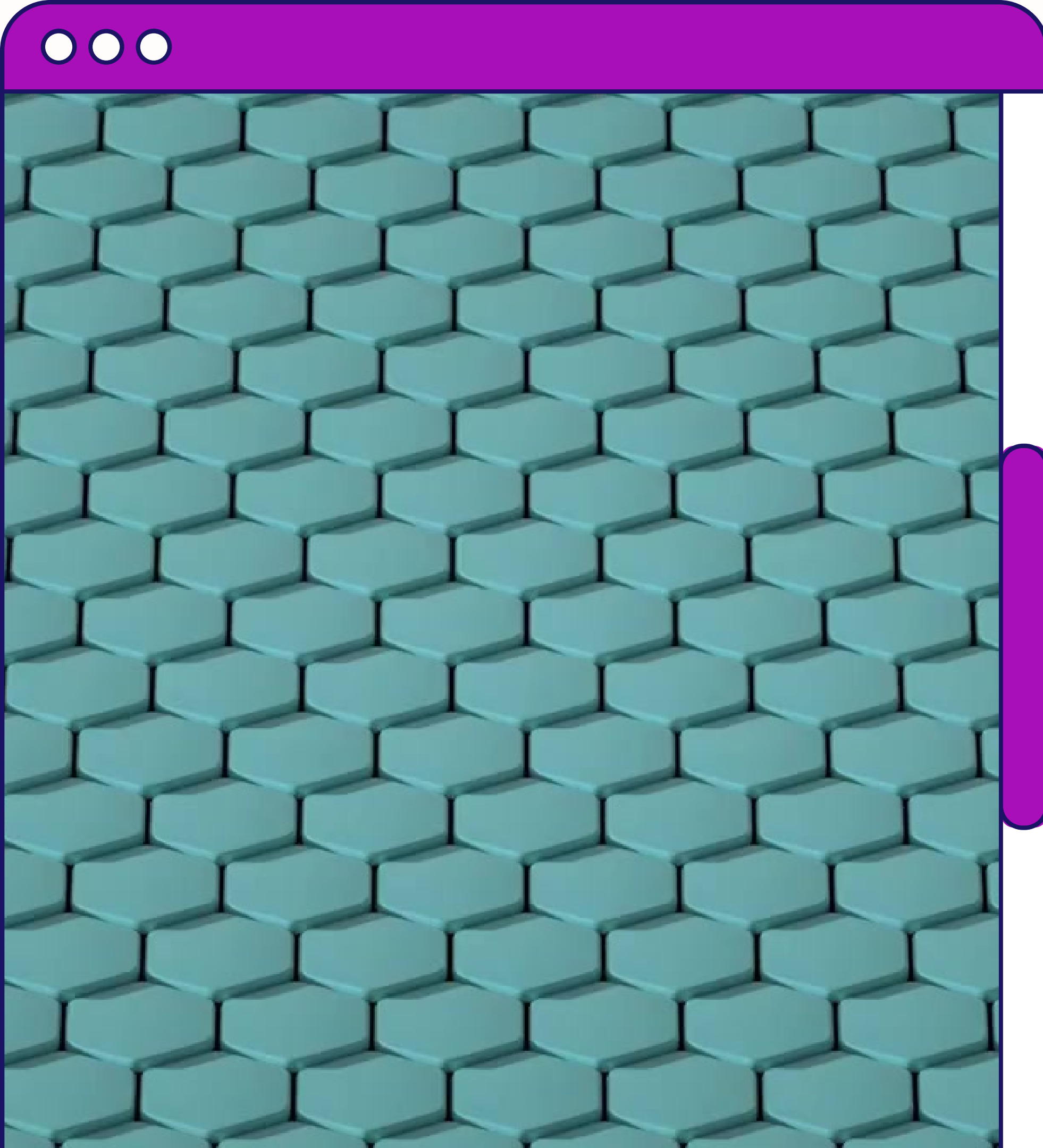
יחודיות



Validity

וולידציה





Consistency

עקביות

Accuracy

Completeness

Consistency

Timeliness

Uniqueness

Validity



איך?



מטרה

2

שדרוג ותוספת
לתהליך קבוע שכבר קיים

4

שימוש חד פעמי
בנתונים חלקיים ע"י שליפה

1

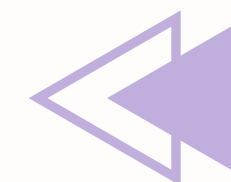
פיתוח חדש
של תהליך קבוע

3

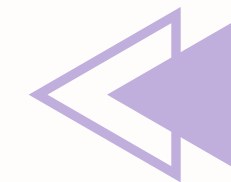
בקרות קבועות
בתהליך קבוע

EVERYONE
HAS A STORY

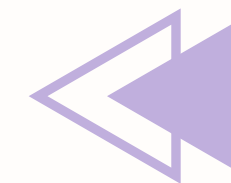
איפה הוא נולד?



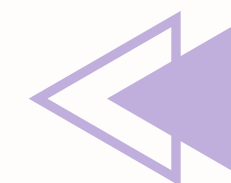
איך הוא גדל?



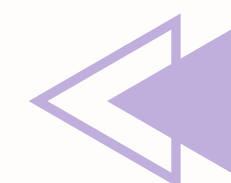
איך הוא נראה היום?



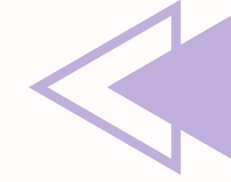
איזה מערכות יחסים יש לו?



יש לו אחים תאומים?



יש לו ילדים?





מיפוי

- 1 תהליכי גזירה והעשרה
- 2 מבנה טבלאות, פורמט שדות
- 3 שדות קשורים ותלויים
- 4 טבלאות אחרות עם מידע זהה
- 5 איזה מערכות המשך ניזונות ממנו ובאיזה אופן



1 מה יקרה במודל אם הערך יהיה שגוי?

2 מה יקרה אם הערך יהיה חסר?

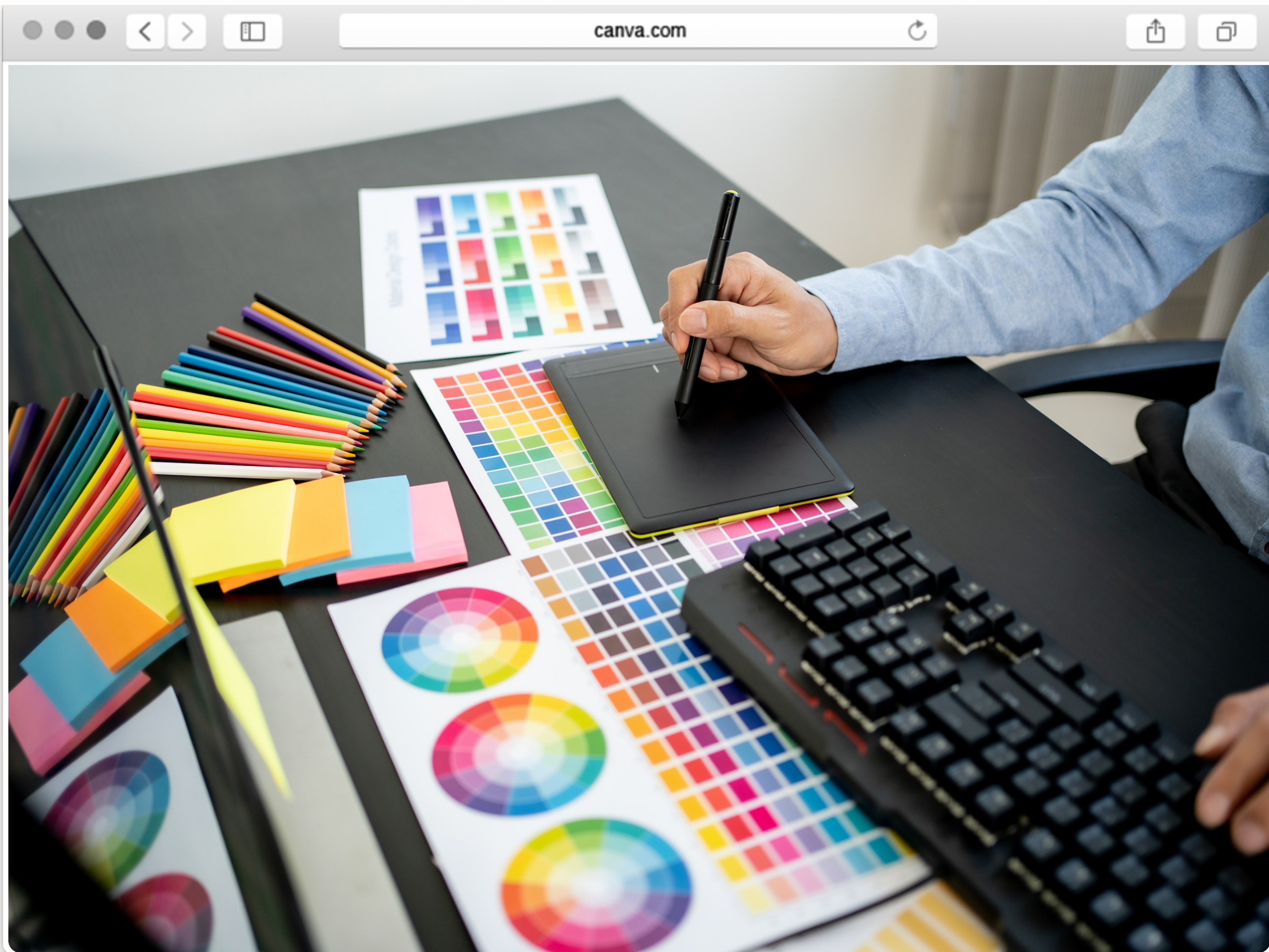
3 האם הערך מגיע ממערכת ליבה/קטלוג מטאדאטה?

4 האם אפשר לבדוק מקבץ שדות יחד?



תכנון היישום

- 1 כלים טכנולוגיים
- 2 תדירות
- 3 מה עושים כשמזהים בעיה (עוצרים? ממשיכים?)
- 4 SLA לטיפול בתקלה
- 5 בדיקות רגרסיה



עיצוב בקרות




- 1 הגדרת מטרות
- 2 מיפוי
- 3 תיעדוף
- 4 תכנון הישום
- 5 עיצוב הבקרות



בקרות שלמות אוכלוסיה

האם יש מקור אחר, זמין ובלתי תלוי? 

מהו תהליך גזירת הנתונים? 

מה נחשב לטעות סבירה מול המקור? 

באיזה רזולוציה בודקים שלמות? 

האם בודקים כמויות או גם סכומים? 



בקרות סבירות ומעקב אחרי מגמות

עומק הבדיקה

3 חודשים? 12 חודשים?



היקף תופעה חריגה בתוך האוכלוסיה

כמות / אחוז



תמהיל לאורך זמן

מוצרים

גורמי סיכון או מאפיינים יחודיים

ותק/שנת חיתום





בקרת איכות

האם הבקרה רלוונטית לערכים?

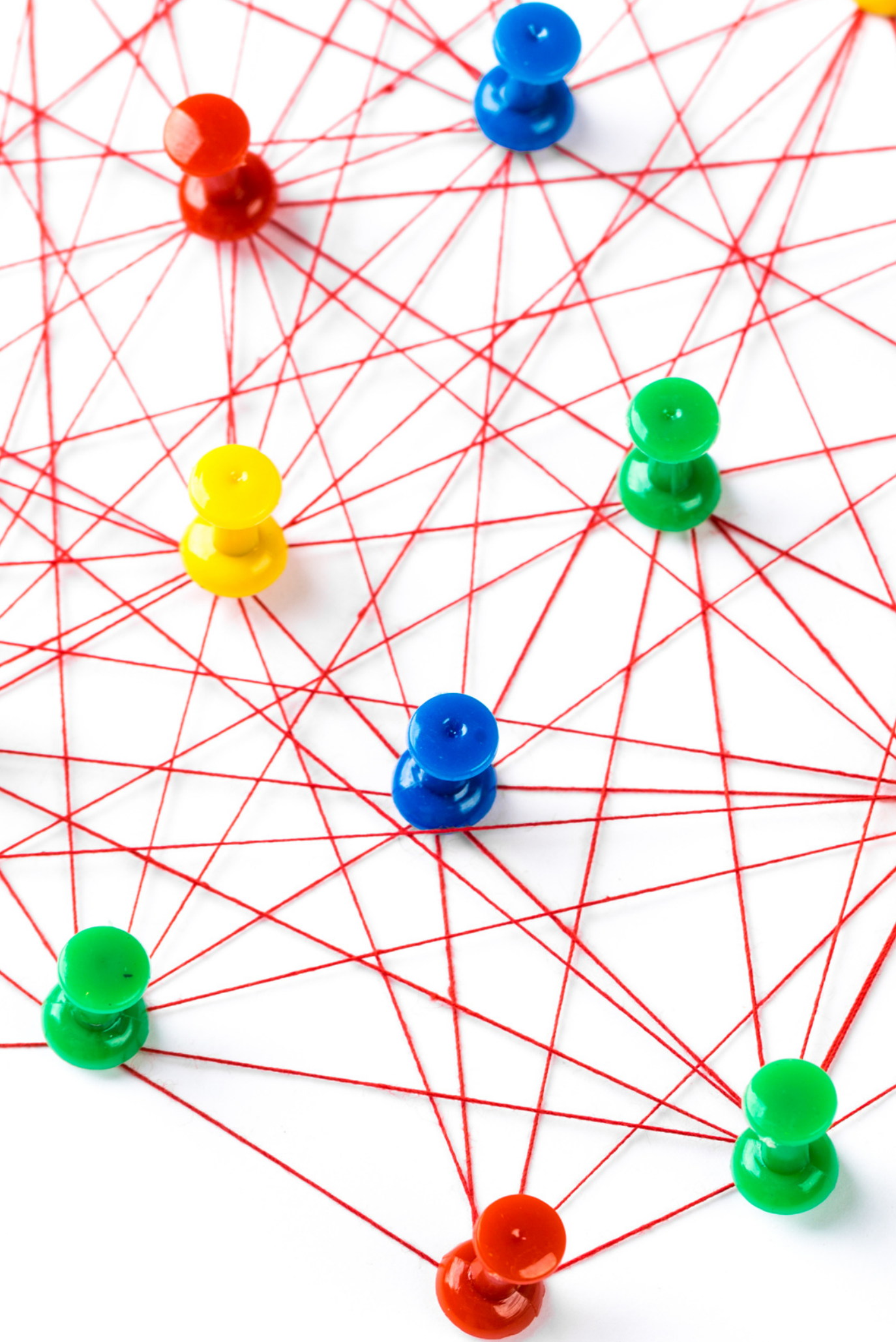
ברירת מחדל נוצרת בתהליך, האם לבדוק ערכים חסרים?

האם זו בקרת ערך חריג או בקרת ערך סביר?

האם לבדוק מינימום ומקסימום? או לבדוק טווח?

האם היא באמת יודעת לזהות תקלות?

האם לבדוק ערך NULL או גם ערך ריק?



בקרת איכות

איך תהליך יצירת נתונים משפיע על הבקרה שלי?

ערך מקטלוג, האם לבדוק מול נתוני מקור בתפעולית?

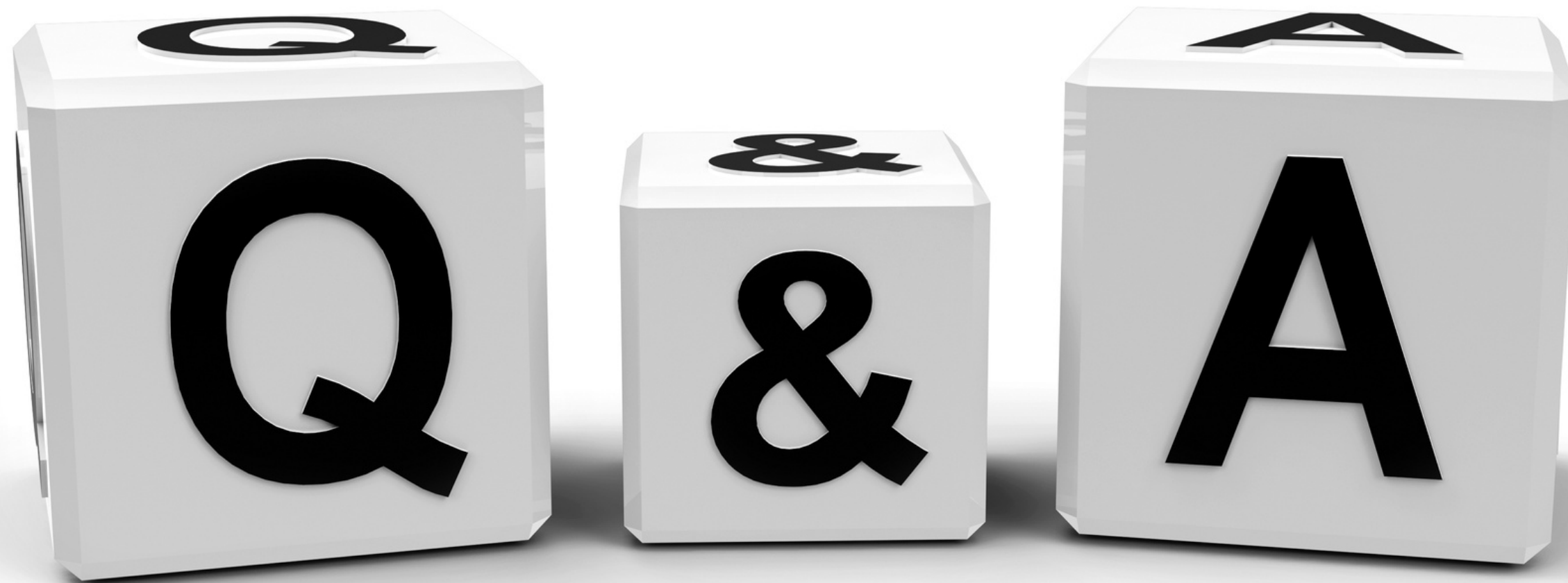
קשרים בין השדות

האם תאריך לידה יכול להיות גדול מתאריך תחילת ביטוח?

האם קיים פרמטר סטטוס עישון לאוכלוסיית ילדים?

האם תאריך סיום יכול להיות יותר נמוך מתאריך תחילה?

האם פרמיה סופית יכולה להיות גבוהה מפרמיה לפני הנחה?



Thanks for listening

