

Εξέταση Μεταβλητότητας

Γιατί Εξηγούμε τη Μεταβλητότητα;

Η εξέταση κατανομών μεμονωμένων μεταβλητών είναι μια καλή αρχή, αλλά το ενδιαφέρον μας είναι **να εξηγήσουμε τη μεταβλητότητα**.

Άτυπος Ορισμός:

Αν γνωρίζουμε την τιμή μιας παρατήρησης σε μια μεταβλητή, μπορούμε να κάνουμε **καλύτερη πρόβλεψη** της τιμής της σε μια άλλη μεταβλητή.

Παράδειγμα:

Αν γνωρίζουμε το **ύψος** ενός ατόμου, μπορούμε να προβλέψουμε καλύτερα το **μήκος του αντίχειρά** του.

Εξαρτημένες vs Ανεξάρτητες Μεταβλητές

Εξαρτημένη Μεταβλητή

Τι είναι: Η μεταβλητή της οποίας τη μεταβλητότητα προσπαθούμε να εξηγήσουμε

Άλλα ονόματα:

- Μεταβλητή αποτελέσματος
- Μεταβλητή απόκρισης

Παράδειγμα: Μήκος αντίχειρα

Ανεξάρτητη Μεταβλητή

Τι είναι: Η μεταβλητή που χρησιμοποιούμε για να εξηγήσουμε τη μεταβλητότητα

Άλλα ονόματα:

- Ερμηνευτική μεταβλητή
- Προβλεπτική μεταβλητή

Παράδειγμα: Ύψος

Λεκτικές Εξισώσεις (Μοντέλα)

Αναπαριστούμε τις σχέσεις με **λεκτικές εξισώσεις**:

μήκος αντίχειρα = ύψος + άλλα πράγματα

Γενική Μορφή:

εξαρτημένη = ανεξάρτητη + άλλα πράγματα

Διάγραμμα Διασποράς

Δημιουργία με R:

```
gf_point(Thumb ~ Height, data = Fingers)
```

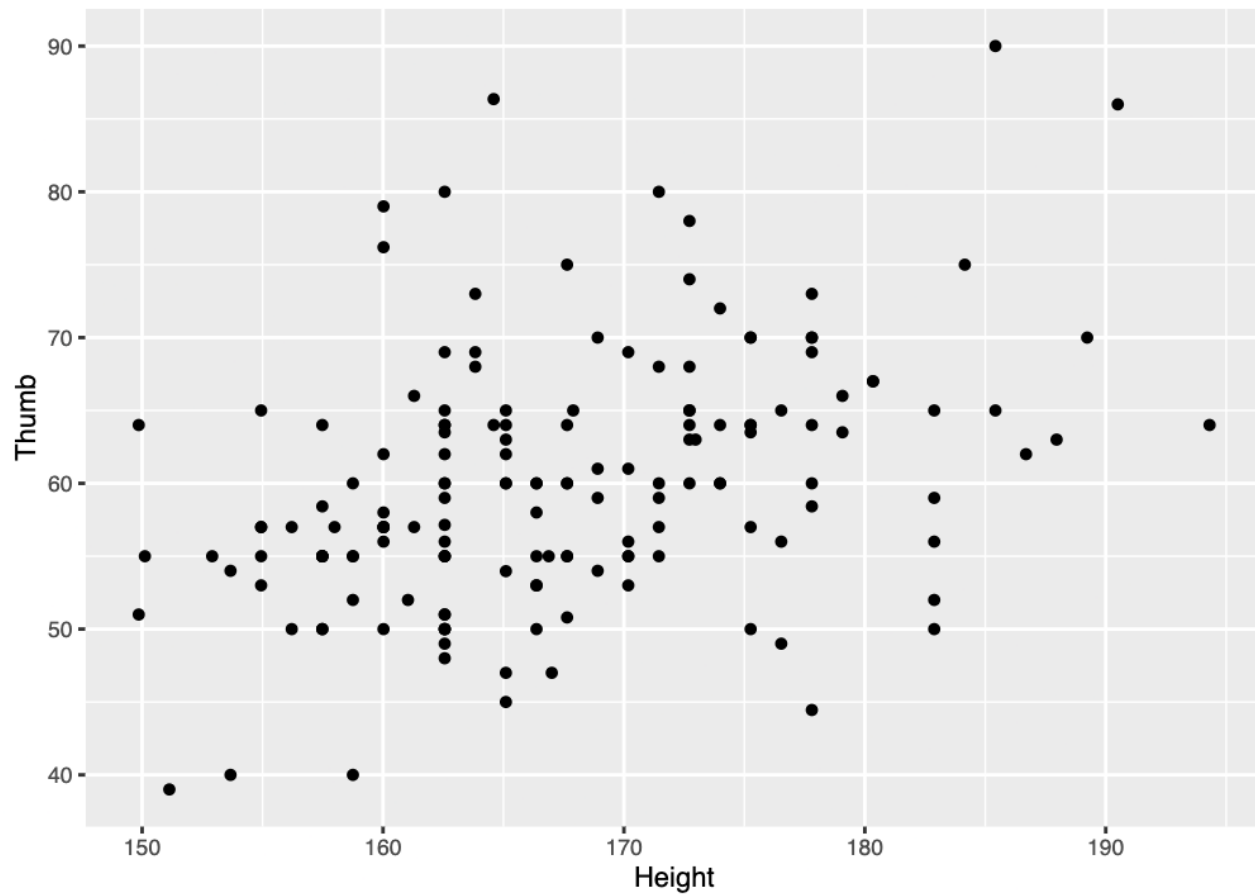
Σύμβαση:

- **Άξονας Y** (κάθετος): Εξαρτημένη μεταβλητή
- **Άξονας X** (οριζόντιος): Ανεξάρτητη μεταβλητή
- Κάθε σημείο = μία παρατήρηση

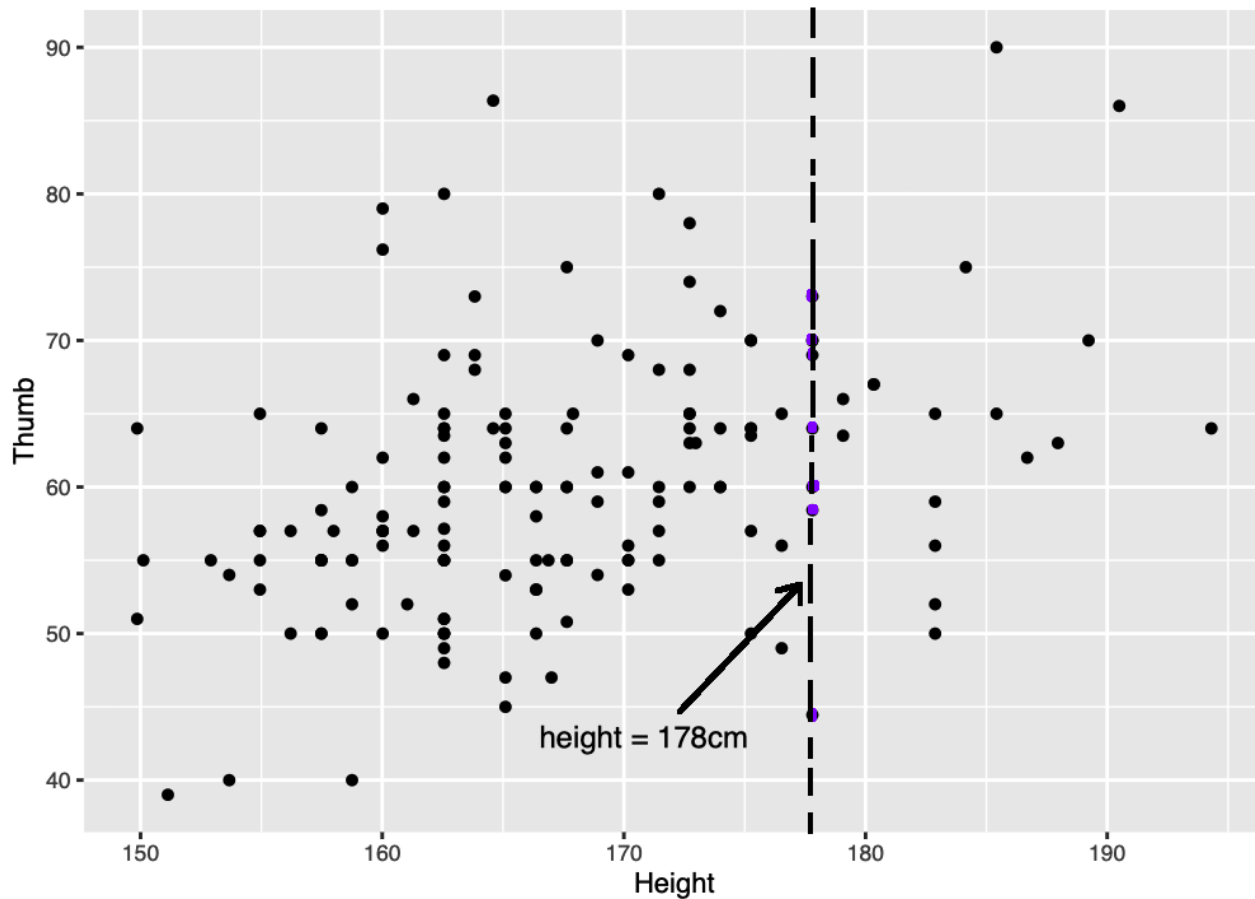
Τι Αναζητούμε:

- Υπάρχει **μοτίβο** ή τάση;
- Καθώς η ανεξάρτητη αυξάνεται, αυξάνεται και η εξαρτημένη;
- Πόσο **διασπορά** υπάρχει γύρω από την τάση;

Διάγραμμα Διασποράς



Διάγραμμα Διασποράς



Ποιοτικές Ανεξάρτητες Μεταβλητές

Όταν η ανεξάρτητη μεταβλητή είναι ποιοτική (π.χ. φύλο):

μήκος αντίχειρα = φύλο + άλλα πράγματα

Διάγραμμα διασποράς με τυχαία μετατόπιση (Jitter plot):

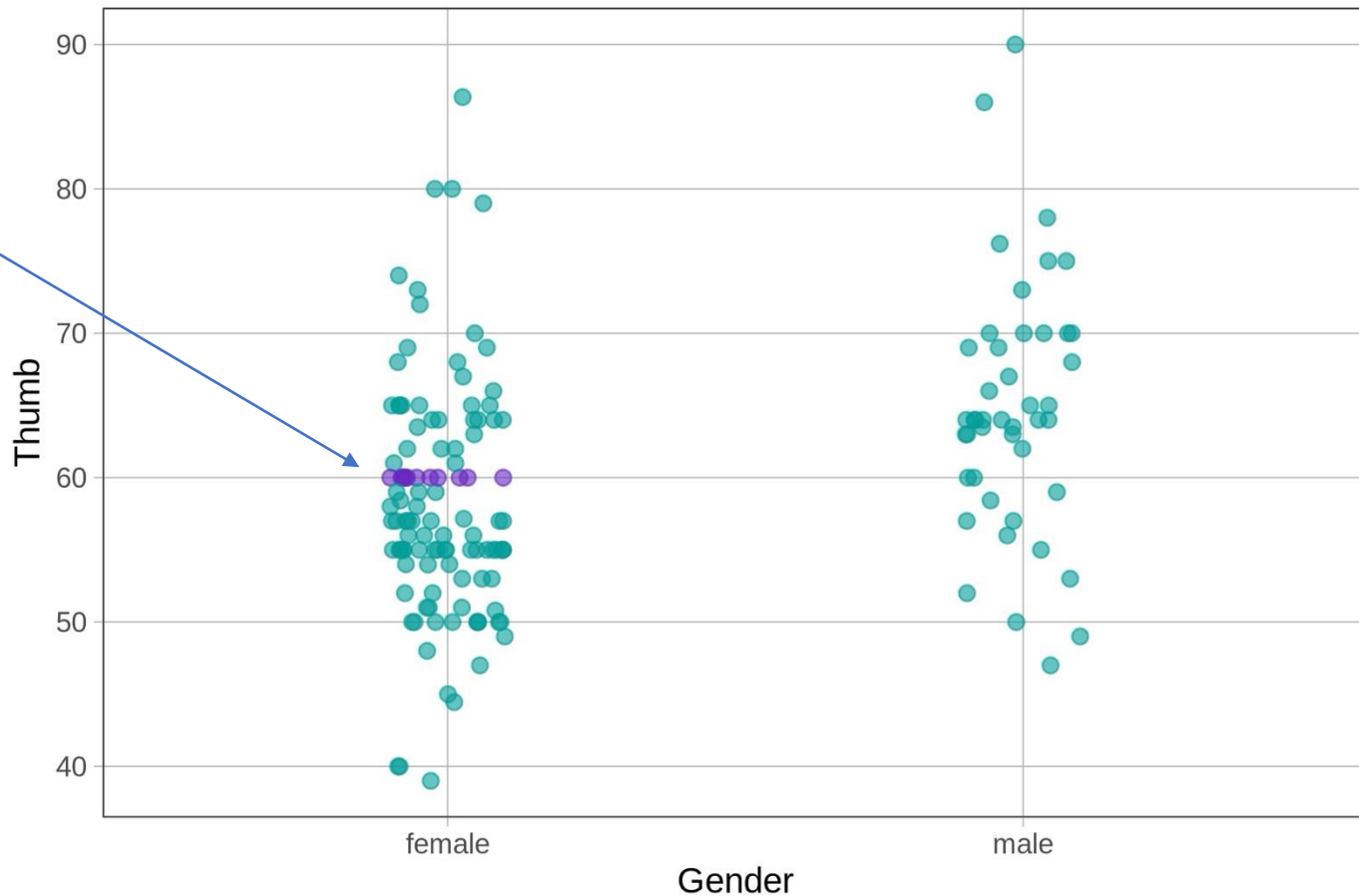
```
gf_jitter(Thumb ~ Gender, data = Fingers)
```

Πλεονεκτήματα:

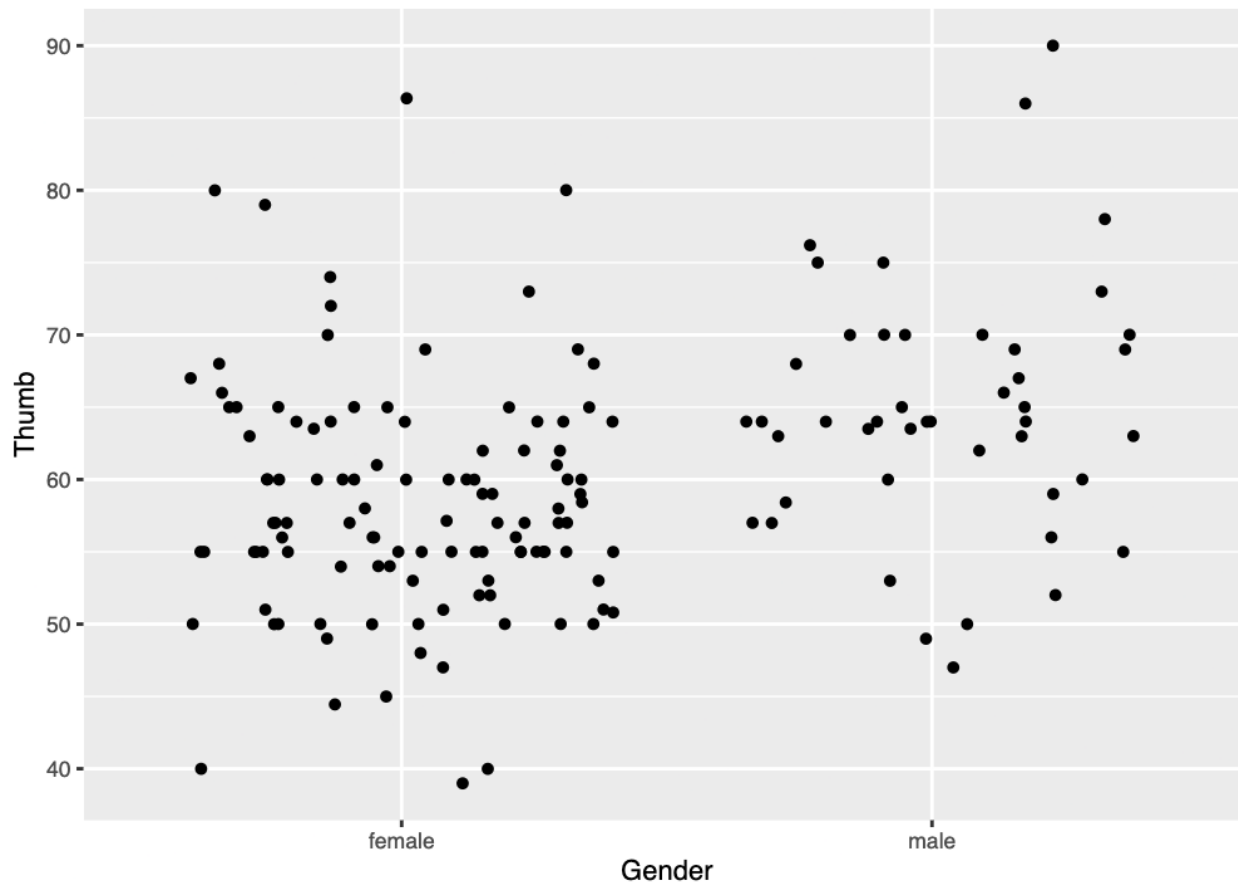
- Αποφεύγει την επικάλυψη σημείων
- Δείχνει όλες τις παρατηρήσεις

Διάγραμμα με τυχαία μετατόπιση (Jitter)

Τα σημεία με
μωβ χρώμα
αντιστοιχούν
στο ίδιο μήκος
αντίχειρα



Διάγραμμα με τυχαία μετατόπιση (Jitter)



Boxplot

Δημιουργία:

```
gf_boxplot(Thumb ~ Gender, data =  
Fingers)
```

Τι Δείχνει το Κουτί:

- Το μεσαίο **50%** των δεδομένων
- Από Q1 (25%) έως Q3 (75%)
- Η παχιά γραμμή = **διάμεσος** (50%)

Ενδοτεταρτημοριακό Εύρος (IQR):

$$\text{IQR} = \text{Q3} - \text{Q1}$$

Απολήξεις:

- Εκτείνονται έως τις ακραίες τιμές
- Όριο: $1.5 \times \text{IQR}$ από το κουτί

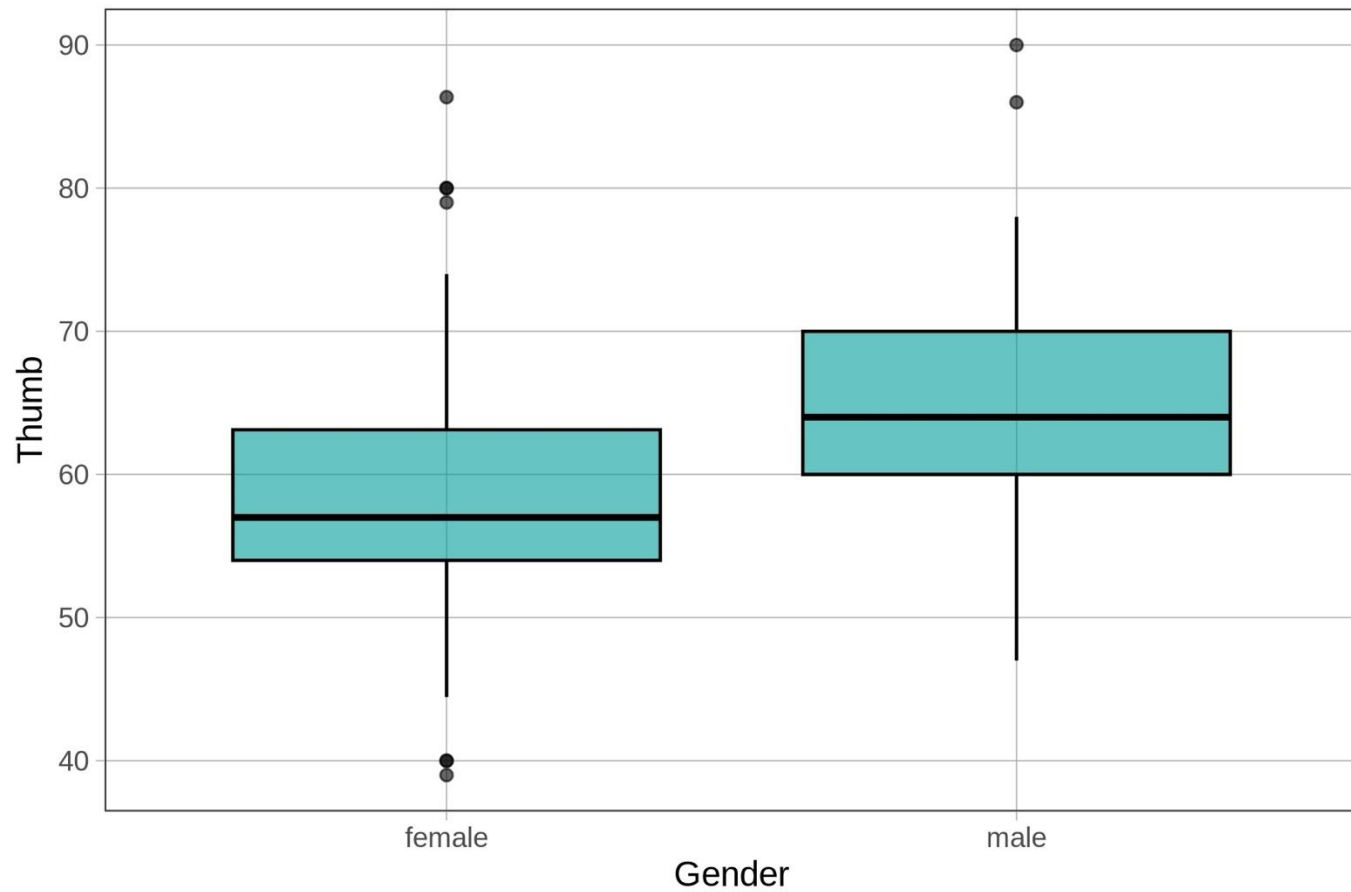
Ακραίες Τιμές:

- Κάτω όριο: $\text{Q1} - 1.5 \times \text{IQR}$
- Πάνω όριο: $\text{Q3} + 1.5 \times \text{IQR}$
- Εμφανίζονται ως μεμονωμένα σημεία

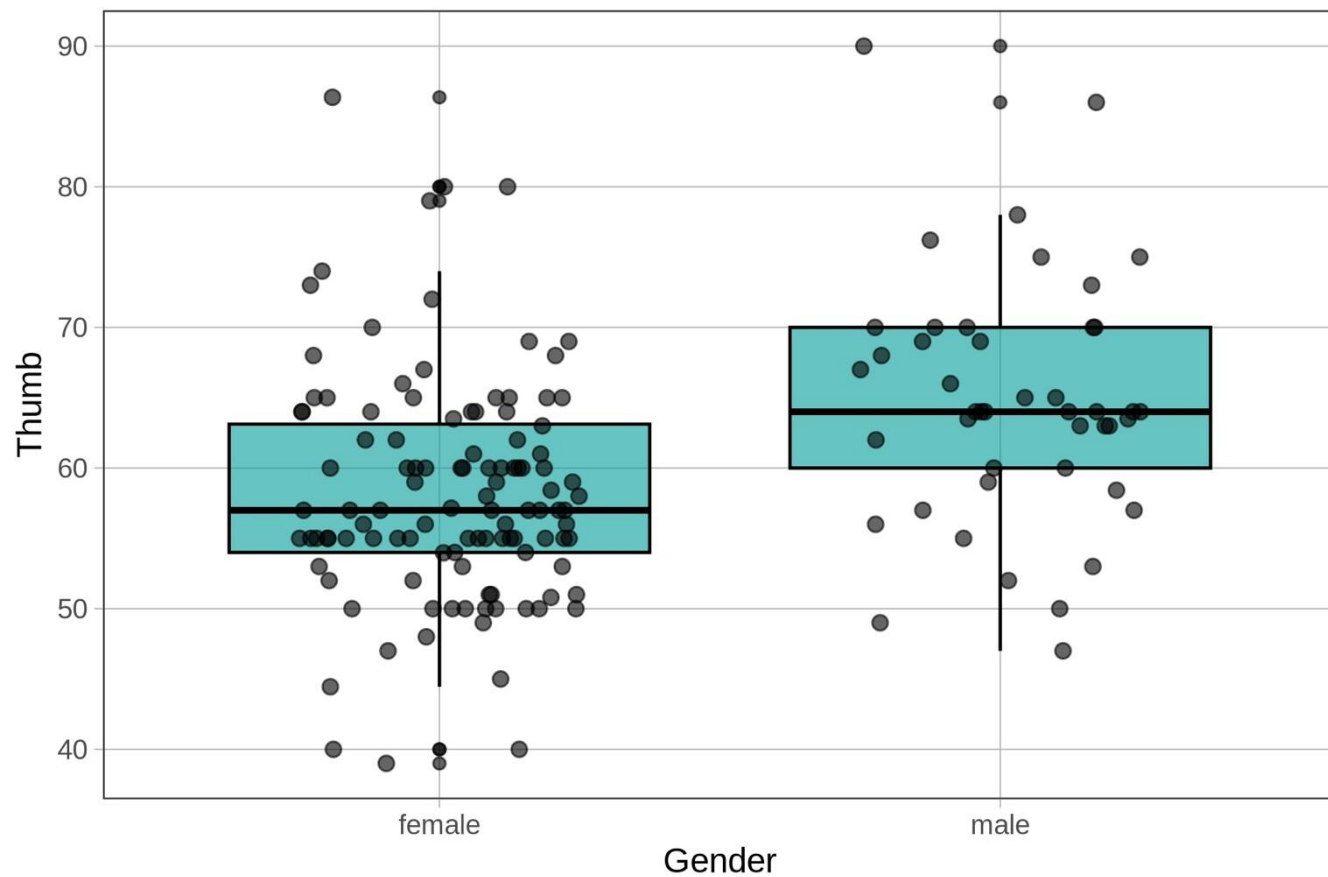
Τι Συγκρίνουμε:

- Κέντρο κατανομών (διάμεσος)
- Διασπορά (εύρος κουτιού)
- Ασυμμετρία

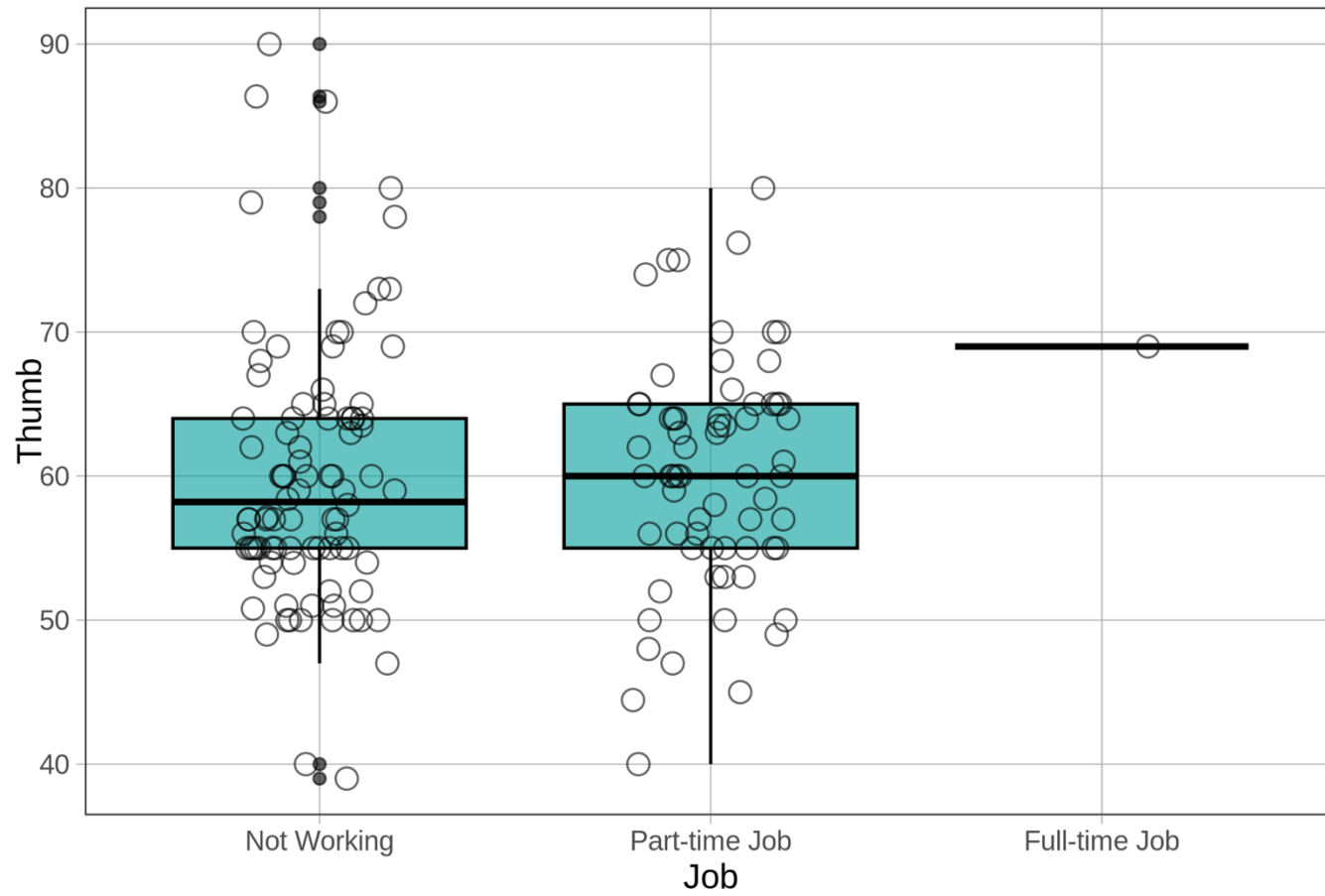
Boxplot



Boxplot + Jitter



Boxplot + Jitter



Διαιρεμένα Ιστογράμματα

Δημιουργία:

```
gf_histogram(~Thumb, data = Fingers) %>%  
  gf_facet_grid(Gender ~ .)
```

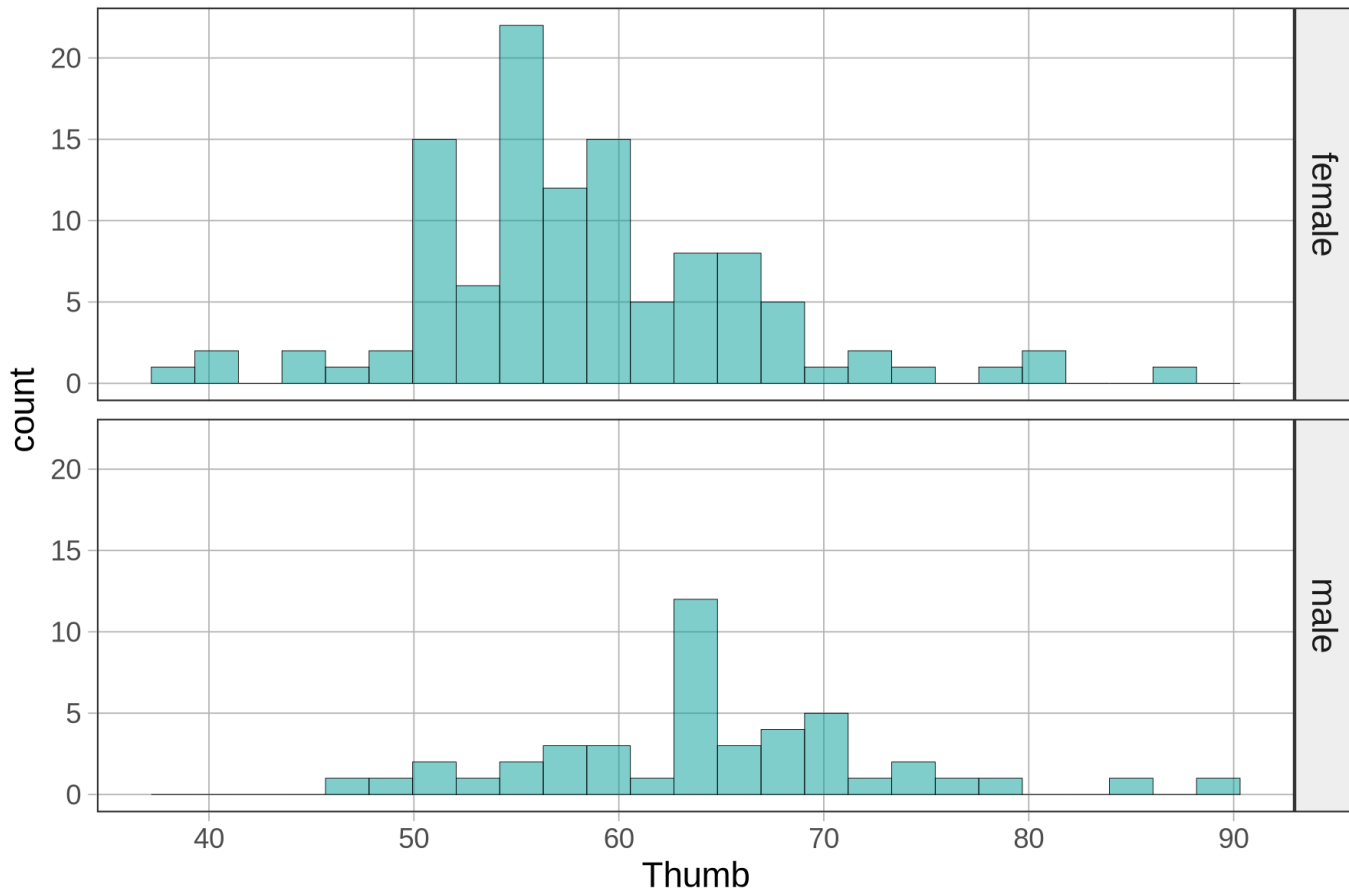
Πυκνότητα Πιθανότητας:

- Χρήσιμη όταν οι ομάδες έχουν διαφορετικό πλήθος
- Επιτρέπει άμεση σύγκριση κατανομών
- Η αναλογία διαιρεμένη με το μήκος διαστήματος (binwidth)

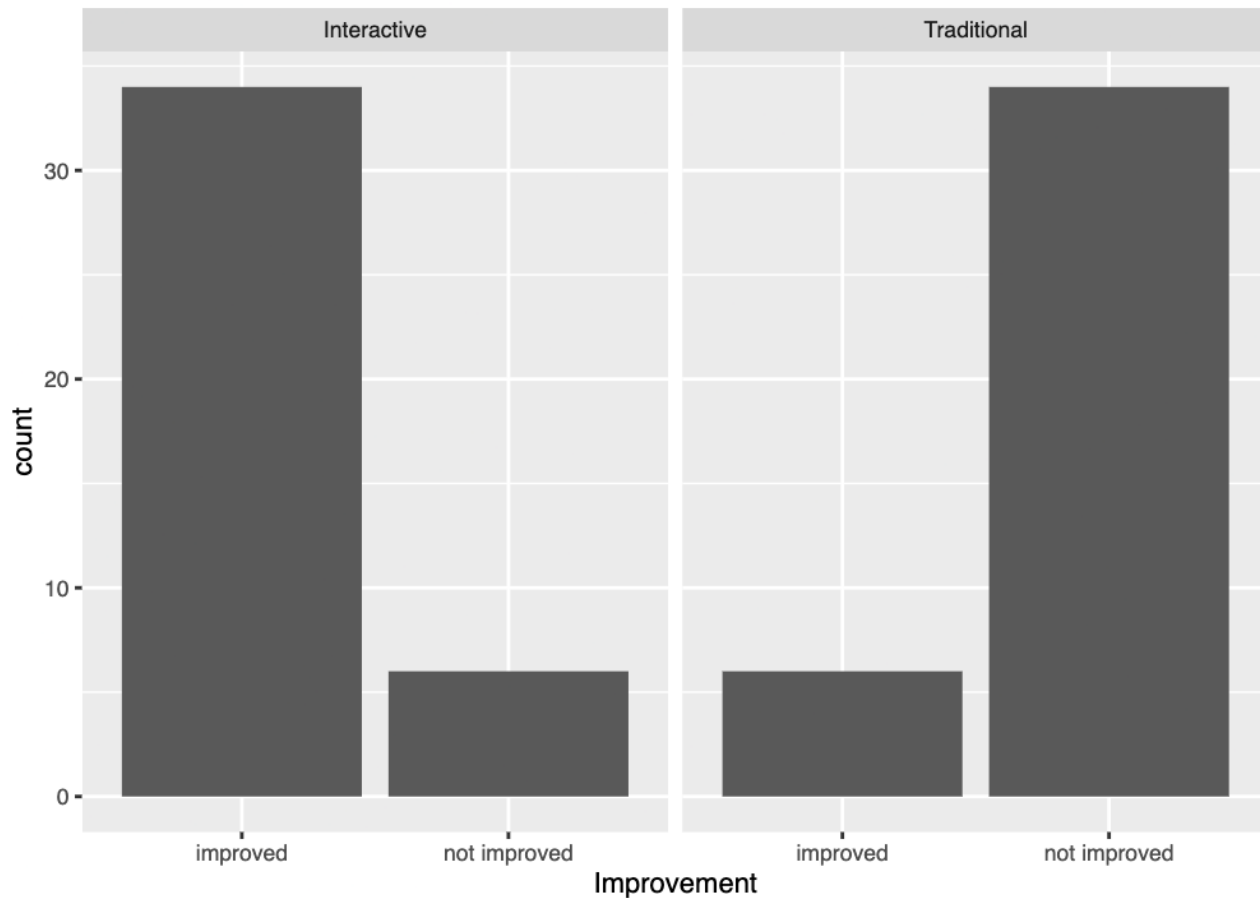
Μεταβλητότητα:

- **Μεταξύ ομάδων:** Διαφορά ανάμεσα στα κέντρα των κατανομών
- **Εντός ομάδας:** Διασπορά μέσα στην ίδια ομάδα

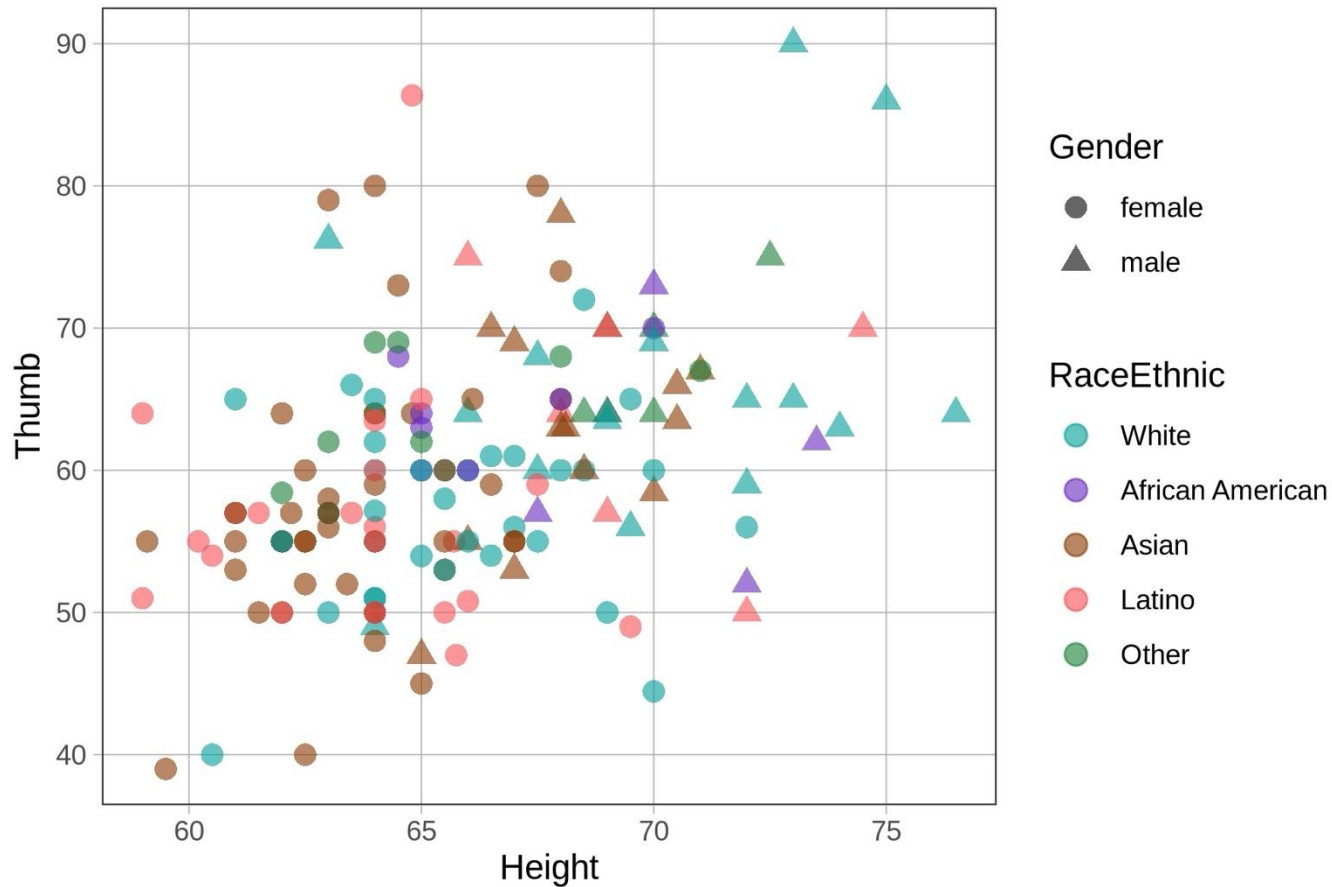
Διαιρεμένα Ιστογράμματα



Διαιρεμένα Ραβδογράμματα



Σύνθετο Διάγραμμα Διασποράς



Βασικές Έννοιες - Ανακεφαλαίωση

1. Εξαρτημένη Μεταβλητή

Αυτή που προσπαθούμε να εξηγήσουμε

2. Ανεξάρτητη Μεταβλητή

Αυτή που χρησιμοποιούμε για εξήγηση

3. Λεκτικές Εξισώσεις

εξαρτημένη = ανεξάρτητη + άλλα πράγματα

4. Οπτικοποιήσεις

Scatterplots, Jitter, Boxplot, Ιστογράμματα

Βασικό Μήνυμα:

Η εξήγηση μεταβλητότητας είναι η καρδιά της στατιστικής ανάλυσης. Χρησιμοποιούμε μια μεταβλητή για να προβλέψουμε ή να κατανοήσουμε μια άλλη.

✓ Σύμβαση R:

Εξαρτημένη ~ Ανεξάρτητη (το ~ διαβάζεται "σε σχέση με")

✓ Άξονες Διαγραμμάτων:

Υ-άξονας = Εξαρτημένη, Χ-άξονας = Ανεξάρτητη

✓ "Άλλα Πράγματα":

Η υπολειπόμενη μεταβλητότητα που δεν εξηγείται

✓ Τελική Πρόβλεψη:

Δεν είναι ποτέ τέλεια, αλλά είναι καλύτερη από το τίποτα!

Ευχαριστώ!

Ερωτήσεις;

Η στατιστική είναι η επιστήμη της εξήγησης
της μεταβλητότητας του κόσμου