



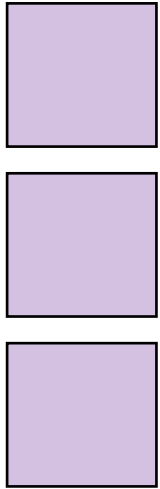
## RDataFrame ve Uproot

$p_x$ [GeV]	$\eta$	$\phi$	$E$ [GeV]	$m$ [GeV]
[1,2,3]				[10,20]
[4,5,6]				[40,50]
[7,8,9]				

Ali Can Canbay  
Ankara Üniversitesi

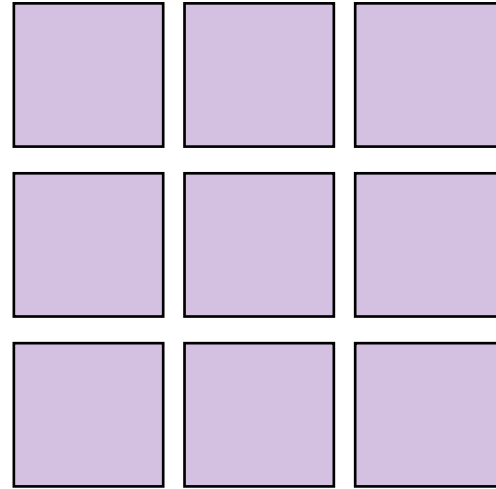
# Vektör – Matris – Veri Çerçevesi

**vektör**



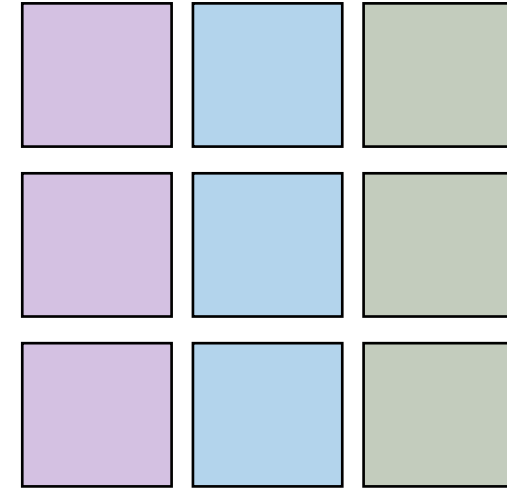
- 1 sütundan oluşur
- Tek veri tipi saklar

**matris**



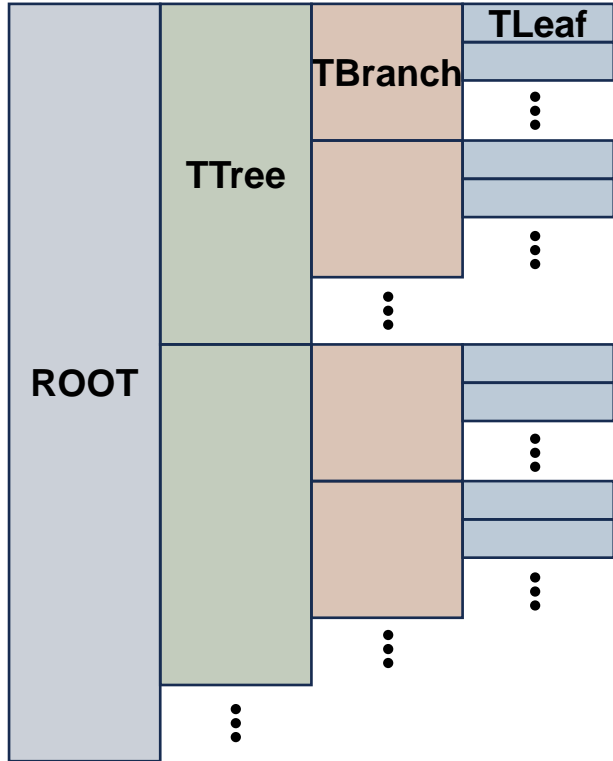
- Çok sayıda satır ve sütundan oluşur
- Tek veri tipi saklar

**veri çerçevesi  
(data-frame)**



- Çok sayıda satır ve sütundan oluşur
- Farklı veri tiplerini saklayabilir

# ROOT Veri Yapısı



**ROOT dosyası** : dosya\_adi.root

**TTree**



: Verilerin ait olduğu ana grupları belirtir:

- Ana veri kümeleri (Delphes, mini, nominal vb.)
- Sistematlere ait veriler (jet sistematiikleri, lepton sistematiikleri vb.)
- Ağırlık verileri (weights)

**TBranch**



: Veri ailelerini belirtir

- Jet
- Lepton

**TLeaf**



: Temel verileri belirtir

- jet\_pt, lepton\_eta vb.

# ROOT Veri Yapısı

$p_t$  : [[58.6, 25.3, ...], [102.7, None, ...], [73.9, 35.2, ...], ...]

$\eta$  : [[1.27, -0.68, ...], [-0.43, None, ...], [1.04, 0.75, ...], ...]

$\phi$  : [[0.7, 1.4, ...], [1.2, None, ...], [0.9, 2.2, ...], ...]

$E$  : [[90, 52, ...], [170, None, ...], [115, 68, ...], ...]

analiz.root														
Delphes														
jet														
Olay	$p_T(j1)$ [GeV]	$p_T(j2)$ [GeV]	...	$\eta(j1)$	$\eta(j2)$	...	$\phi(j1)$	$\phi(j2)$	...	$E(j1)$ [GeV]	$E(j2)$ [GeV]	...	...	...
1	58.6	25.3		1.27	-0.68		0.7	1.4		90	52		...	...
2	102.7	–		-0.43	–		1.2	–		170	–		...	...
3	73.9	35.2		1.04	0.75		0.9	2.2		115	68		...	...
...														

## Temel verilere nasıl erişilir?

C++ ve Python için hazırlanmış olan temel ROOT ve pyROOT kütüphaneleri kullanıldığında:

- Olay döngüsü oluşturulur
  - Parçacık döngüsü oluşturulur ve şartlar (kesme, cut) test edilir.
  - Hesaplamalar yapılır.
  - Sonuçlar döngü dışı nesnelere eklenir
- Sonuçlar elde edilir.

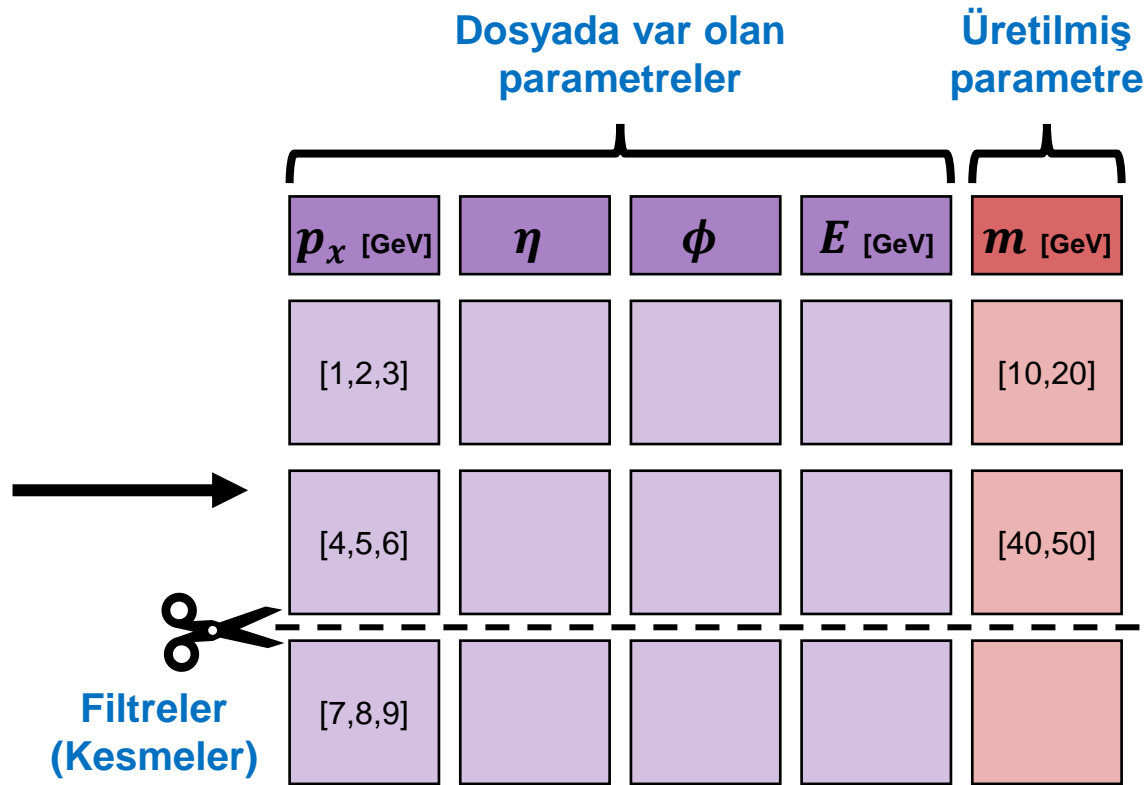
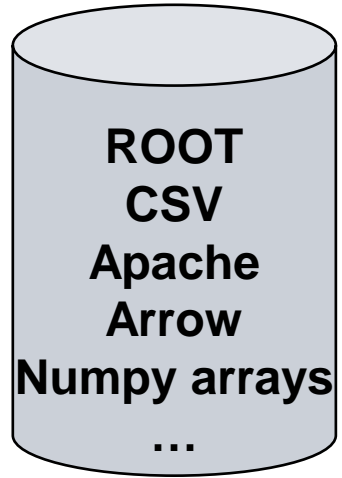
Bu durumda, olay döngüsünün her bir adımında işlemler tekrarlanır ve nispeten (özellikle Python için) yavaş ve karışık bir yöntemdir.

ROOT Data Frame (RDataFrame - RDF), ROOT'un resmi analiz araçlarından biridir (Python modülü). Veriler veri çerçevesi nesnesi (data-frame object) olarak tanımlanır.

$p_x$ [GeV]	$\eta$	$\phi$	$E$ [GeV]
[1,2,3]			
[4,5,6]			
[7,8,9]			

- pandas vb. veri çerçevelerinde alışılmış olan yapıdadır.
- ROOT ve pyROOT'un izin verdiği işlemler tanımlıdır.
- Parçalı işlemler ve paralelleştirme işlemleri kolaydır.

Veri Dosyası



Histogramlar

Yeni ROOT dosyası

Kesme akışı (cut flow) raporları

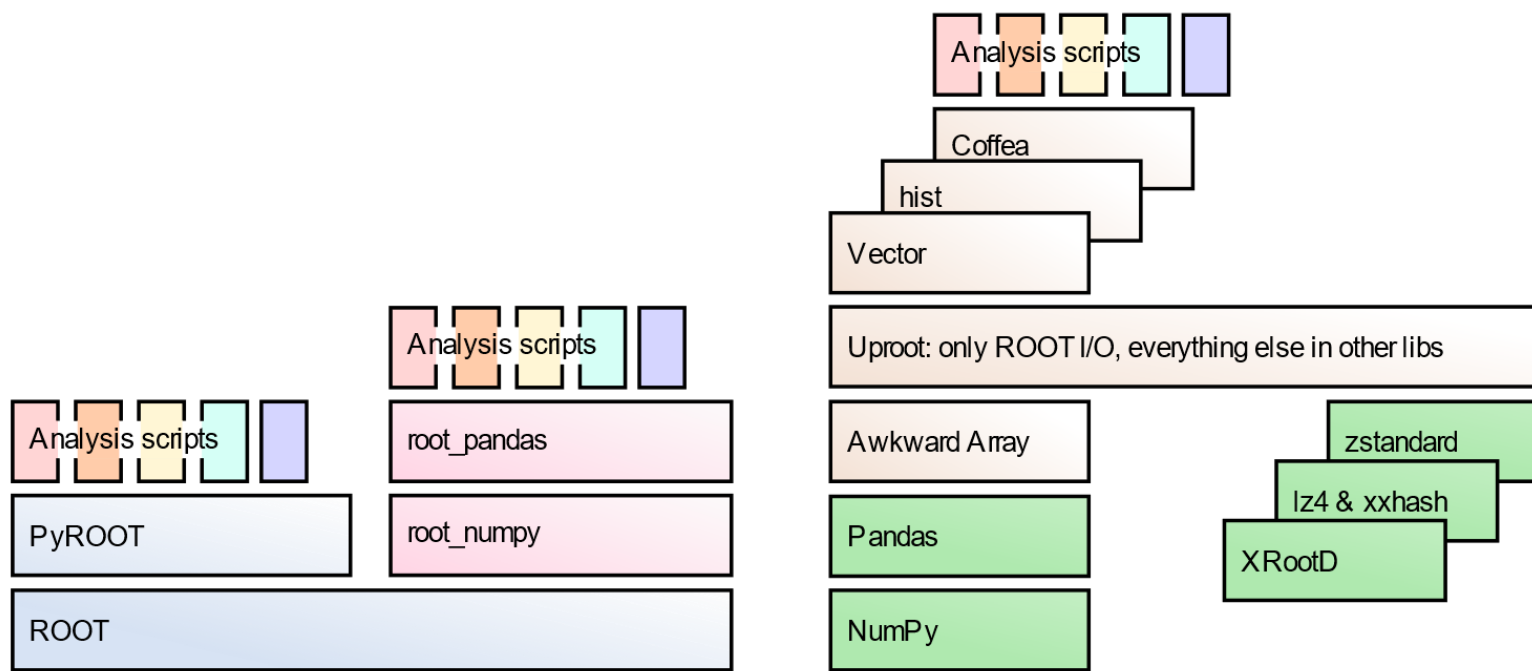
Veri bilgileri (ortalama, std, ...)

Kullanıcı tarafından tanımlanmış diğer işlemler

# Uproot

Uproot, ROOT yazılımından bağımsız olarak yüklenip kullanılabilir. Arkaplanda C++ kütüphaneleri kullanmaz (Python modülü).

- ROOT dosyaları içerisindeki verileri numpy dökümleri olarak alır.
- Verileri Python'un makine öğrenmesi modüllerine direkt aktarılacak şekilde saklar.
- Kullanımı, Python sözlük yapıları veya JSON verileri ile benzerlik gösterir.



<https://uproot.readthedocs.io/en/latest/index.html>



1.  $pp \rightarrow H X \rightarrow b \bar{b} X$  sürecine ait olayları barındıran `ornek_Hbb.root` dosyasındaki verileri kullanarak Higgs kütlelerini RDataFrame kullanarak yeniden yapılandırınız.
2.  $pp \rightarrow Z \rightarrow l^+ l^-$  sürecine ait olayları barındıran `ornek_Zll.root` dosyasındaki verileri kullanarak Z-bozon kütlelerini Uproot kullanarak yeniden yapılandırınız.

## **RDataFrame Kaynakları**

[pyROOT](#)

[ROOT Tutorials](#)

[RDataFrame Manual](#)

[RDataFrame Tutorials](#)

## **Uproot Kaynakları**

[Uproot](#)

[Awkward Array](#)

[Vector](#)

[Matplotlib](#)

[Scikit-HEP Tutorial](#)

[Uproot Tutorial](#)

# Ek: Histogramlar

Bir veri kümesine ait frekans dağılımlarını gösteren grafiklere histogram adı verilir.

Ölçüm No	Kütle [gr]	Ölçüm No	Kütle [gr]
1	110	16	98
2	95	17	112
3	102	18	108
4	118	19	100
5	101	20	93
6	96	21	98
7	103	22	109
8	102	23	92
9	98	24	105
10	119	25	99
11	83	26	101
12	104	27	104
13	96	28	93
14	92	29	88
15	101	30	105

**Bu elmaların kütlelerinin histogramını çizebilmek için:**

- Histogramdaki kutu genişliğinin (bin size) belirlenmesi
- Bu genişliğin frekansının belirlenmesi gerekir.

## 5 gr kutu genişliği

Aralık [gr]	Frekans
<80	0
80-84	1
85-89	1
90-94	4
95-99	7
100-104	9
105-109	4
110-114	2
115-119	2
>120	0

## 10 gr kutu genişliği

Aralık [gr]	Frekans
<80	0
80-89	2
90-99	11
100-109	13
110-119	4
>120	0

# Ek: Histogramlar

