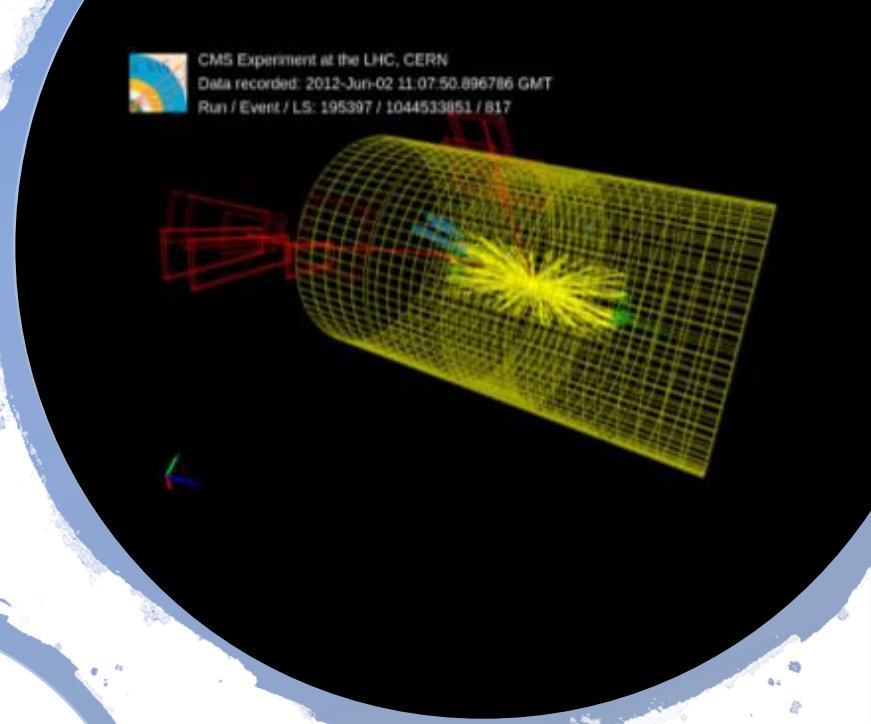
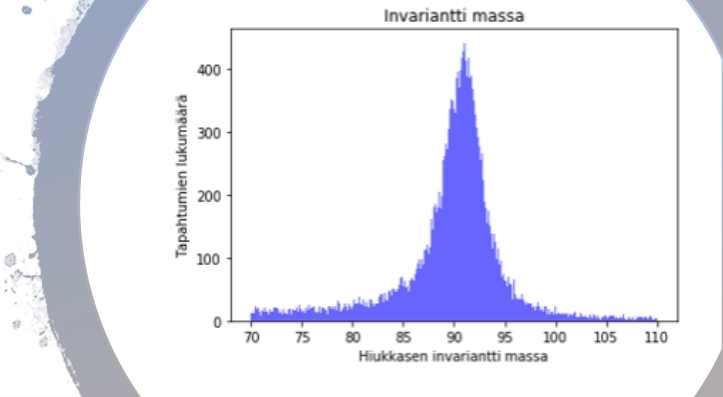

TANELI
EERO
CERN
opendata
TEAM
JENNY

What was our job?

- Test CERN's the new material for students
- Practise data analysis: for example, how to draw histograms and calculate invariant mass
- Produce own data by doing experiments (pendulum movement)
- Learn how to use Jupyter notebook and Python 3.0
- Increase our knowledge from programming and data concerning particle detectors



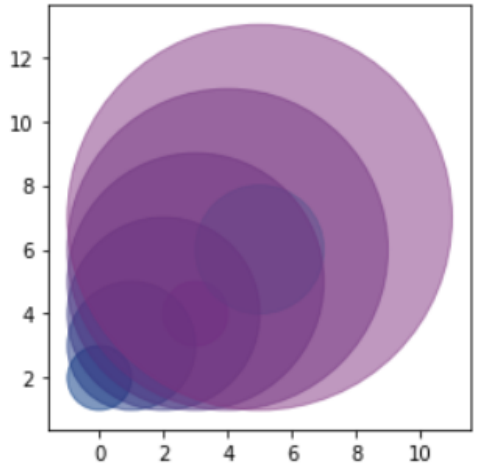


In [1]:
```
import matplotlib.pyplot as plt
```

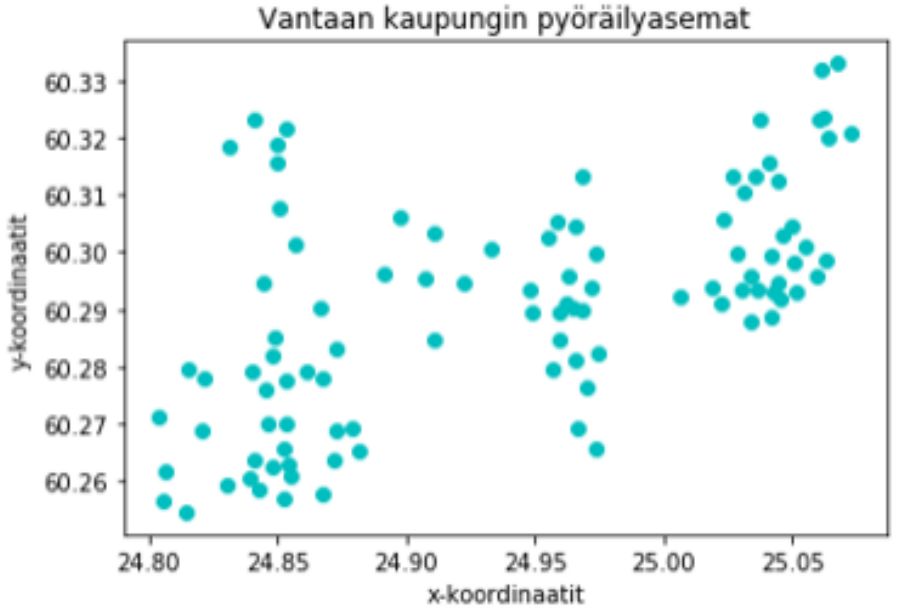
In [2]:
```
def create_circle(x,c,a,b):
    circle=plt.Circle((a,b), radius=x, color=c, alpha=0.5)
    return circle
```

In [3]:
```
def show_shape(patch):
    ax=plt.gca()
    ax.add_patch(patch)
    plt.axis('scaled')
```

In [6]:
```
c=create_circle(1,"m",3,4)
show_shape(c)
d=create_circle(2,"c",5,6)
show_shape(d)
for x in range(6):
    show_shape(create_circle(x+1,(x*0.1, 0.2, 0.5),x,x+2))
plt.show()
```

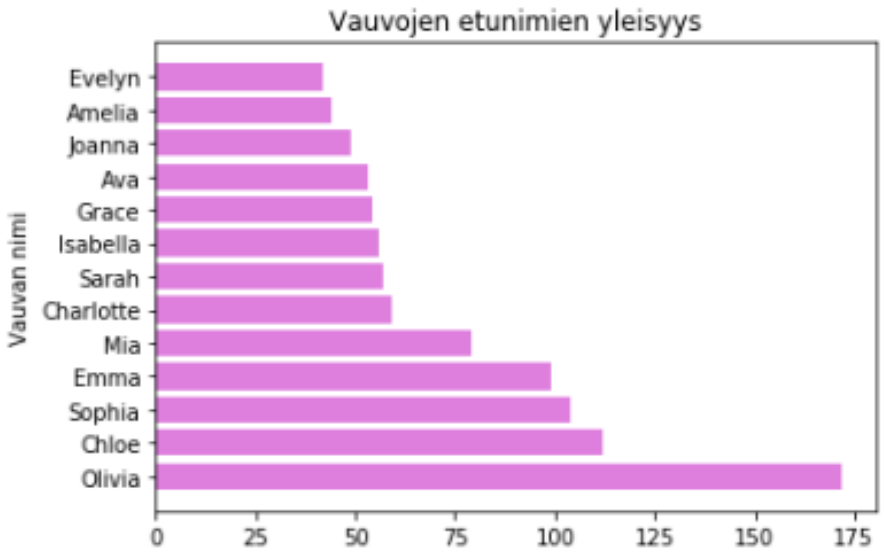



```
import pandas as pd
import numpy as np
import matplotlib.pylab as plt

d=pd.read_csv("asemat.csv")

d.head()

plt.plot(d.x , d.y, "o", color="c")
plt.title("Vantaan kaupungin pyöräilyasemat")
plt.xlabel("x-koordinaatit")
plt.ylabel("y-koordinaatit")
plt.show()
```



	Year of Birth	Gender	Ethnicity	nimi	lkm	Rank
0	2016	FEMALE	ASIAN AND PACIFIC ISLANDER	Olivia	172	1
1	2016	FEMALE	ASIAN AND PACIFIC ISLANDER	Chloe	112	2
2	2016	FEMALE	ASIAN AND PACIFIC ISLANDER	Sophia	104	3
3	2016	FEMALE	ASIAN AND PACIFIC ISLANDER	Emma	99	4
4	2016	FEMALE	ASIAN AND PACIFIC ISLANDER	Mia	79	5

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
d=pd.read_csv("vauvat.csv")

x=d.nimi
y=d.lkm
plt.barh(x, y, alpha=0.5, color="m")

plt.ylabel("Vauvan nimi")
plt.xlabel("Nimien lukumäärä")
plt.title("Vauvojen etunimien yleisyys")
plt.show()
```

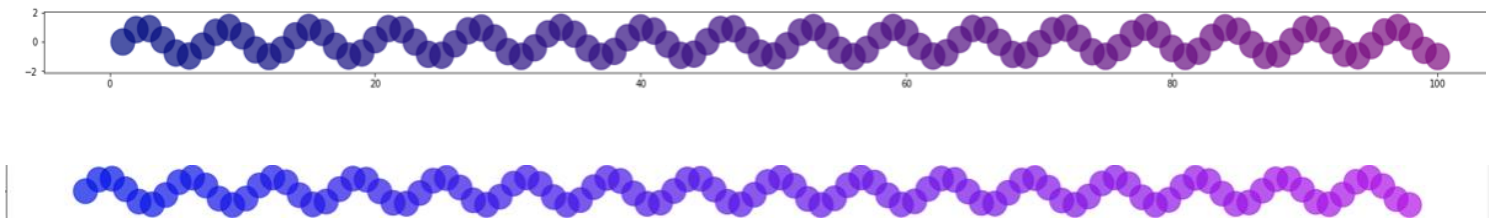


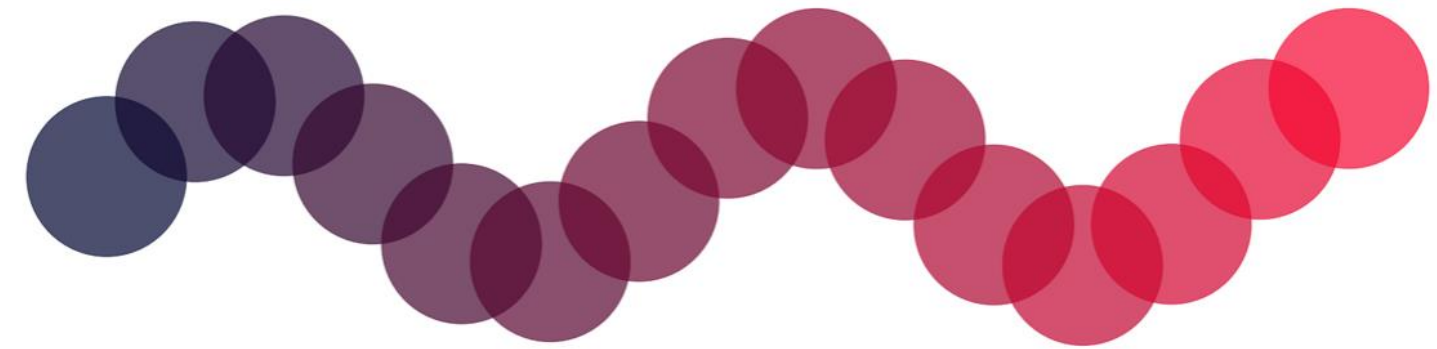

```
import matplotlib.pyplot as plt
import numpy as np
```

```
def create_circle(x,c,a,b):
    circle=plt.Circle((a,b), radius=x, color=c, alpha=0.7)
    return circle
```

```
def show_shape(patch):
    ax=plt.gca()
    ax.add_patch(patch)
    plt.axis('scaled')
```

```
plt.figure(figsize = (30,250))
c=create_circle(0.2, "m", 10,10)
#show_shape(c)
#d=create_circle(0.5,"c",50,50)
#show_shape(d)
for x in range(100):
    show_shape(create_circle(0.9,(x*0.005, 0.05, 0.5),x+1, np.sin(x)))

plt.show()
```



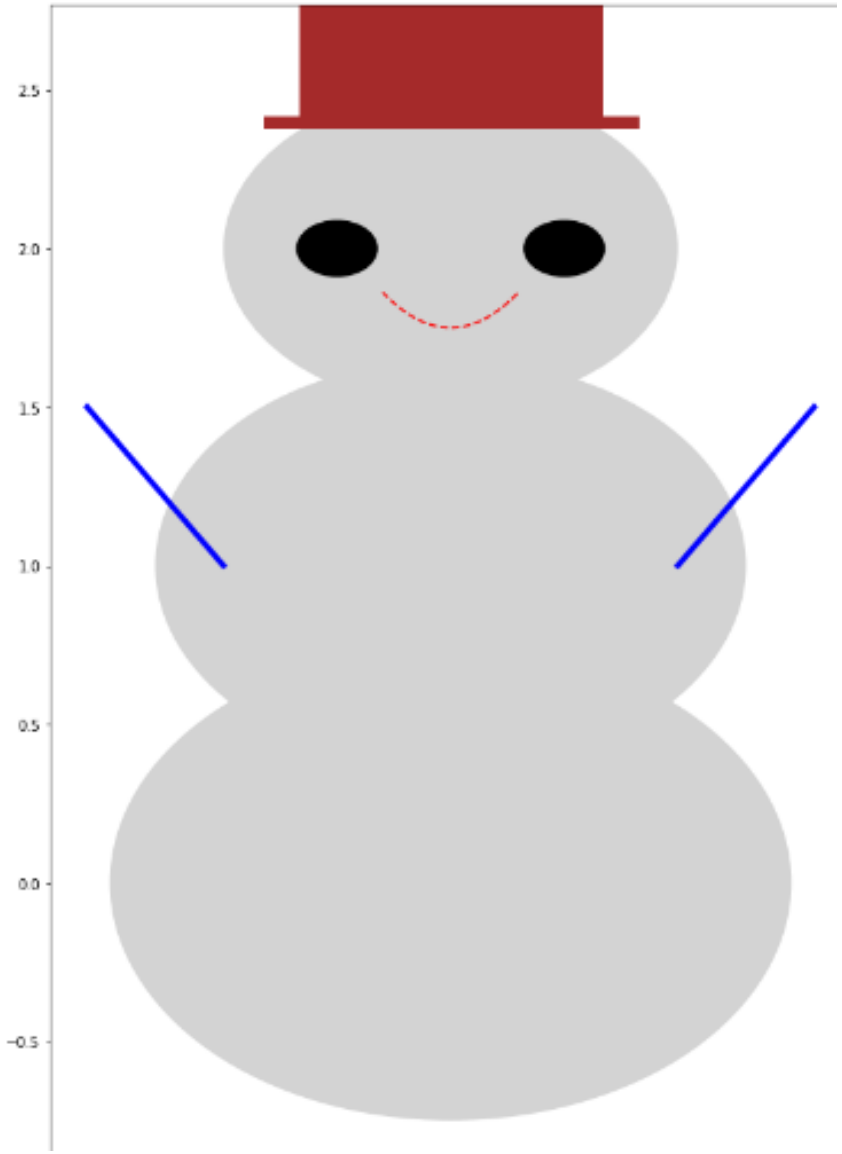


Olen iloinen

a 5.00

b 1.00

c 1.00

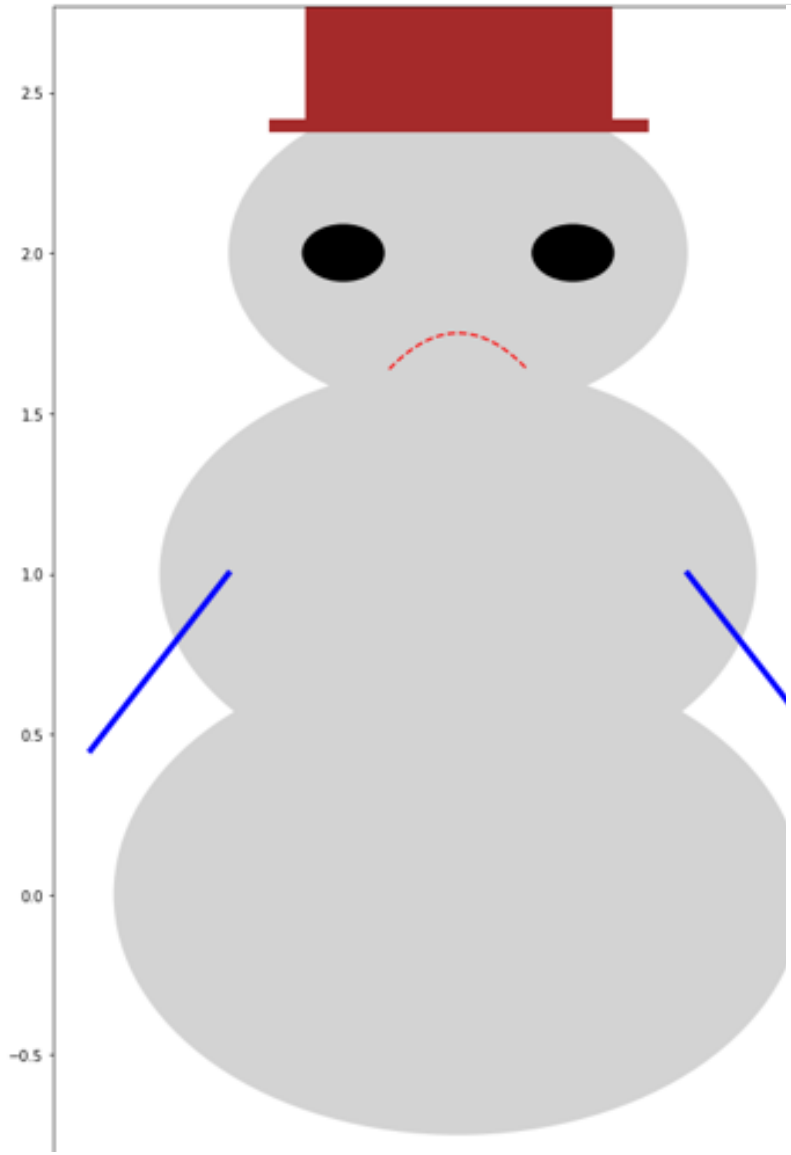


Olen surullinen

a -5.00

b 0.30

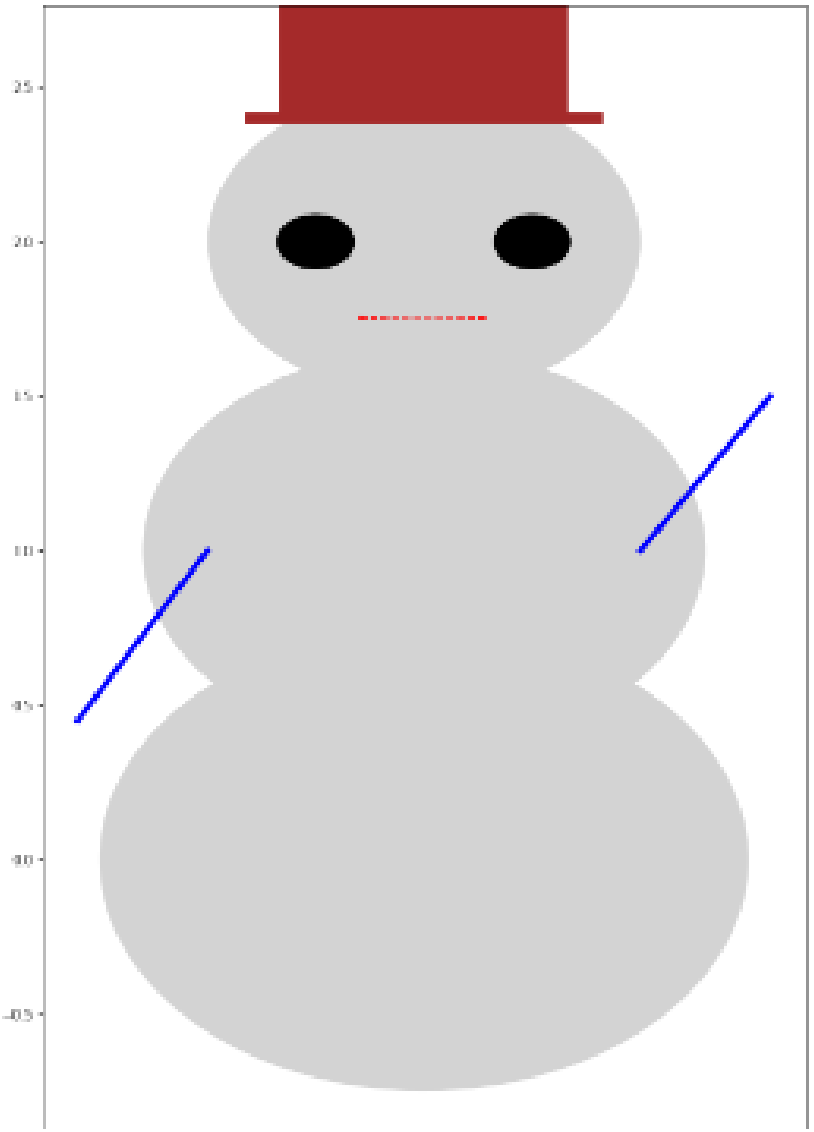
c 0.30



a 0.20

b 0.30

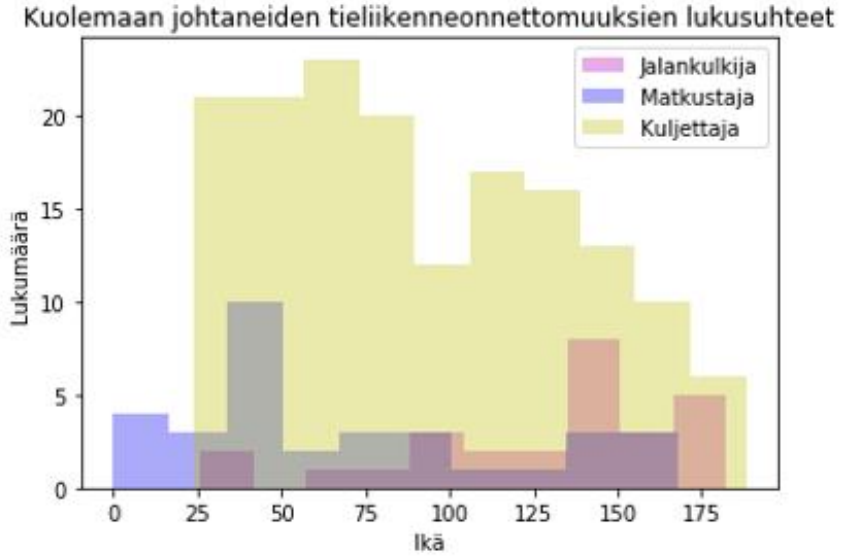
c 1.00



HUONO VERSIO!!

In [223]:
```
JK=Kuollut+Jalankulkija
KU=Kuollut+Kuljettaja
MA=Matkustaja+Kuollut
```

In [224]:
```
plt.hist(JK.Ikä, alpha=0.33, label=("Jalankulkija"), color="m")
plt.hist(MA.Ikä, alpha=0.33, label=("Matkustaja"), color="b")
plt.hist(KU.Ikä, alpha=0.33, label=("Kuljettaja"), color="y")
plt.legend()
plt.title("Kuolemaan johtaneiden tieliikenneonnettomuuksien lukusuhteet")
plt.xlabel("Ikä")
plt.ylabel("Lukumäärä")
plt.show()
```

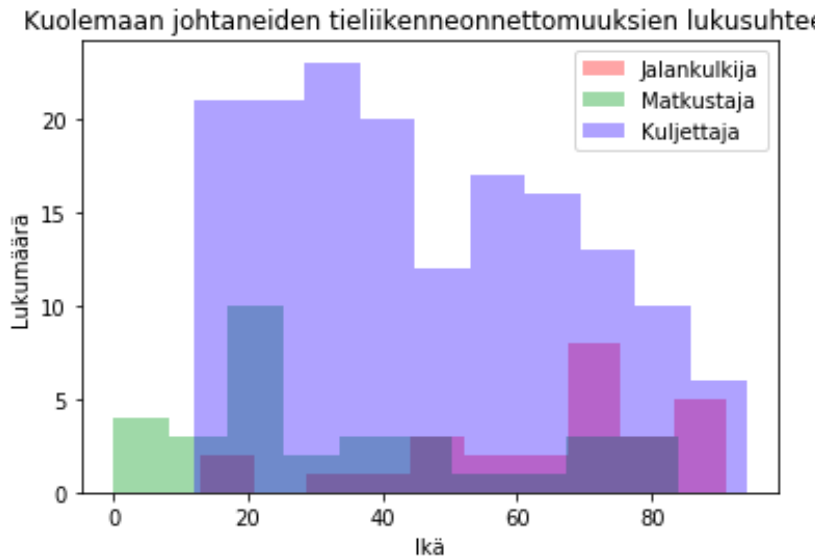


Kuolemaan johtaneiden onnettomuuksien lukusuhteet

HYVÄ VERSIO!

In [217]:
```
KJK=pd.merge(Kuollut,Jalankulkija)
KMA=pd.merge(Kuollut,Matkustaja)
KKU=pd.merge(Kuollut,Kuljettaja)
```

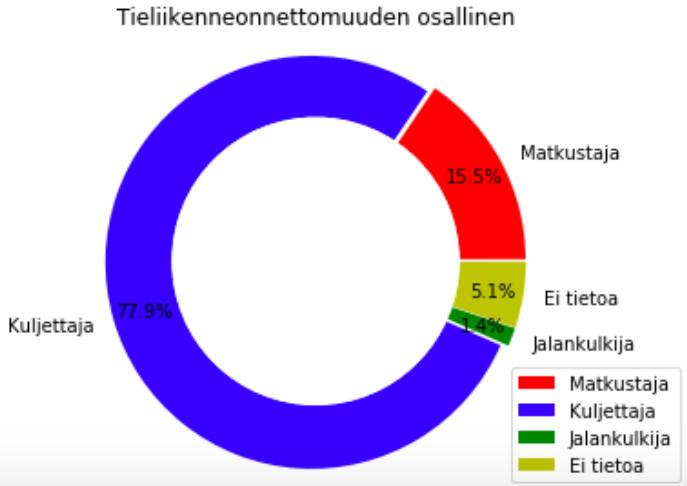
In [218]:
```
plt.hist(KJK.Ikä, alpha=0.33, label=("Jalankulkija"), color="r")
plt.hist(KMA.Ikä, alpha=0.33, label=("Matkustaja"), color="g")
plt.hist(KKU.Ikä, alpha=0.33, label=("Kuljettaja"), color="b")
plt.legend()
plt.title("Kuolemaan johtaneiden tieliikenneonnettomuuksien lukusuhteet")
plt.xlabel("Ikä")
plt.ylabel("Lukumäärä")
plt.show()
```



Tieliikenneonnettomuudessa ollut osallinen

In [215]:
```
Kuljettaja=data0[(data0["Kulj_matk"]=="KU")]
Matkustaja=data0[(data0["Kulj_matk"]=="MA")]
Jalankulkija=data0[(data0["Kulj_matk"]=="JK")]
Ei_dataa=data0[(data0["Kulj_matk"]=="XX")]
```

In [216]:
```
labels=['Matkustaja','Kuljettaja', 'Jalankulkija', 'Ei tietoa']
sizes=[len(Matkustaja)/len(data0), len(Kuljettaja)/len(data0), len(Jalankulkija)/len(data0), len(Ei_dataa)/len(data0)]
colors=['r', 'b', 'g', 'y']
explode = (0.02, 0.02, 0.02, 0.02)
plt.pie(sizes, labels=labels, explode=explode, colors=colors,autopct="%1.1f%%", pctdistance=0.85)
plt.title("Tieliikenneonnettomuuden osallinen")
plt.axis('equal')
centre_circle = plt.Circle((0,0),0.70,fc='white')
fig = plt.gcf()
fig.gca().add_artist(centre_circle)
plt.tight_layout()
plt.legend()
plt.show()
```

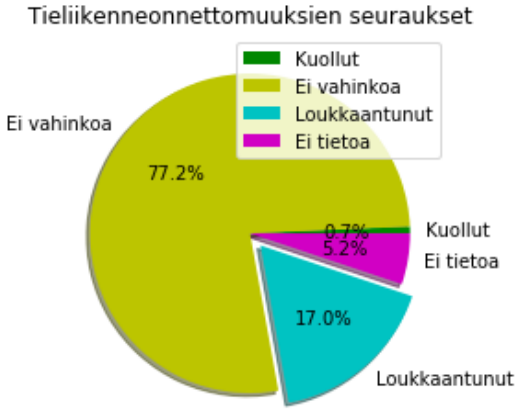


Tieliikenneonnettomuuksien seuraukset

In [212]:
```
data0 = data.replace(np.NaN,0)
```

In [213]:
```
Kuollut = data0[(data0["Seuraussel"]=="kuollut")]
Ei_vahinkoa = data0[(data0["Seuraussel"]=="ei henkilövahinkoa")]
Loukkaantunut = data0[(data0["Seuraussel"]=="loukkaantunut")]
Ei_tietoa = data0[(data0["Seuraussel"]==0)]
```

In [214]:
```
labels=['Kuollut', 'Ei vahinkoa', 'Loukkaantunut', 'Ei tietoa']
sizes=[len(Kuollut)/len(data0), len(Ei_vahinkoa)/len(data0), len(Loukkaantunut)/len(data0), len(Ei_tietoa)/len(data0),]
colors=['g', 'y', 'c', 'm']
explode = (0, 0, 0.1, 0)
plt.pie(sizes, labels=labels, explode=explode, colors=colors,autopct='%1.1f%%', shadow=True)
plt.title("Tieliikenneonnettomuuksien seuraukset")
plt.legend()
plt.show()
```

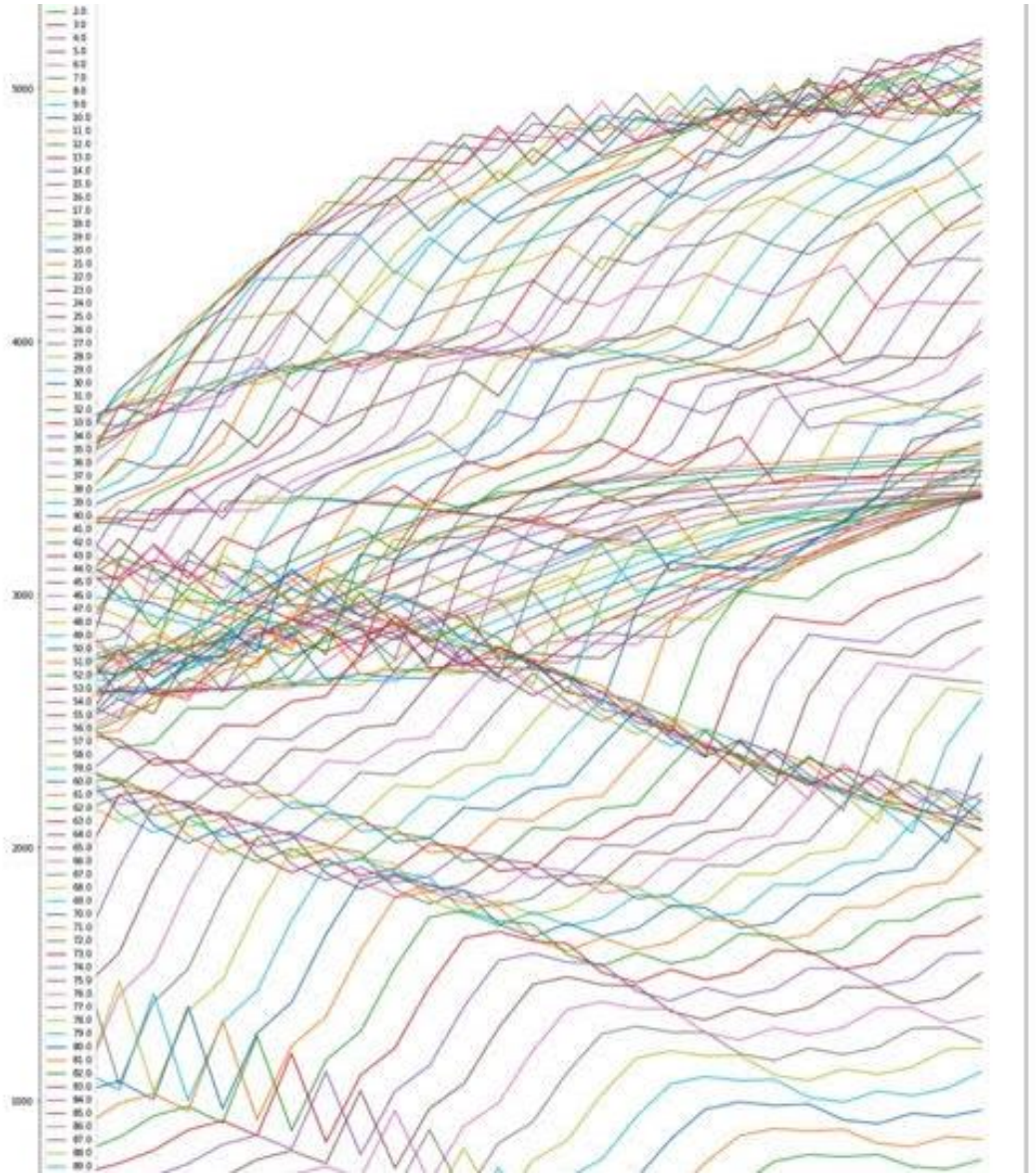


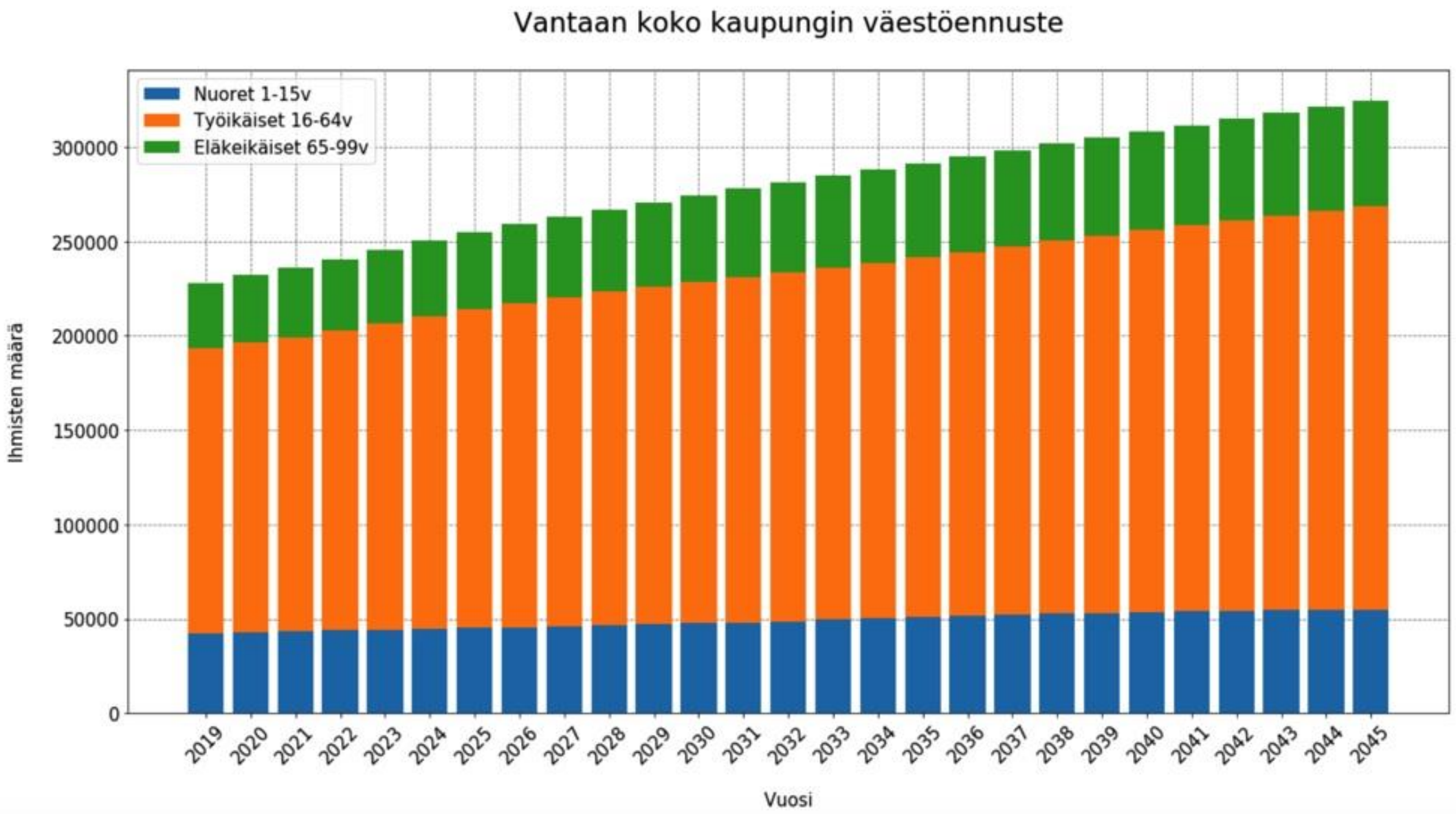
Esimerkki aineistosta

	Sukupuoli	Age	2019	2020	2021	2022	2023	2024	2025	2026	...	2036	2037	2038	2039	2040	2041	2042	2043	2044
1	Kaikki	0.0	2576.0	2687.0	2775.0	2850.0	2928.0	3010.0	3084.0	3152.0	...	3469.0	3481.0	3490.0	3499.0	3507.0	3516.0	3524.0	3534.0	3544.0
2	Kaikki	1.0	2613.0	2622.0	2720.0	2823.0	2899.0	2981.0	3055.0	3125.0	...	3486.0	3497.0	3509.0	3519.0	3528.0	3536.0	3546.0	3554.0	3563.0
3	Kaikki	2.0	2622.0	2618.0	2614.0	2725.0	2828.0	2906.0	2979.0	3049.0	...	3442.0	3457.0	3468.0	3480.0	3489.0	3498.0	3505.0	3513.0	3521.0
4	Kaikki	3.0	2593.0	2624.0	2610.0	2616.0	2726.0	2831.0	2902.0	2971.0	...	3399.0	3415.0	3429.0	3440.0	3451.0	3459.0	3468.0	3474.0	3483.0
5	Kaikki	4.0	2779.0	2603.0	2625.0	2619.0	2626.0	2738.0	2836.0	2903.0	...	3368.0	3385.0	3400.0	3414.0	3424.0	3435.0	3443.0	3451.0	3457.0
6	Kaikki	5.0	2722.0	2784.0	2599.0	2630.0	2624.0	2632.0	2737.0	2832.0	...	3329.0	3348.0	3364.0	3379.0	3393.0	3402.0	3413.0	3420.0	3428.0
7	Kaikki	6.0	2666.0	2727.0	2782.0	2604.0	2634.0	2629.0	2632.0	2735.0	...	3280.0	3313.0	3331.0	3347.0	3361.0	3375.0	3384.0	3394.0	3401.0
8	Kaikki	7.0	2719.0	2681.0	2737.0	2797.0	2619.0	2651.0	2642.0	2643.0	...	3246.0	3282.0	3314.0	3333.0	3349.0	3363.0	3376.0	3385.0	3395.0
9	Kaikki	8.0	2780.0	2727.0	2683.0	2744.0	2804.0	2628.0	2655.0	2645.0	...	3197.0	3238.0	3274.0	3306.0	3324.0	3339.0	3353.0	3366.0	3375.0

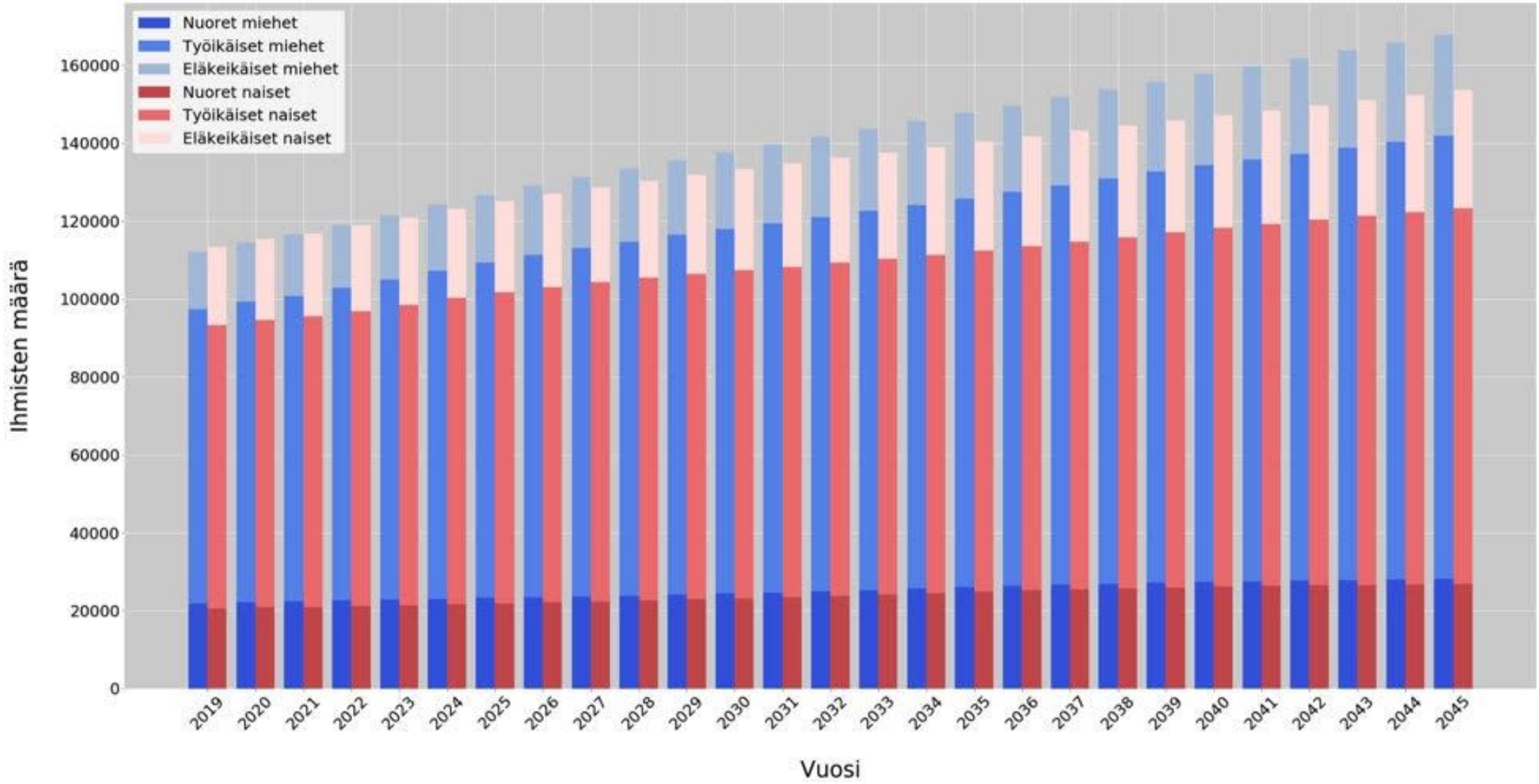
```
fig = plt.figure(figsize=(40, 20), edgecolor='white')
ax = fig.add_subplot(1,1,1)
ax.set_axisbelow(True)
ax.set_facecolor('lightgray')
ax.yaxis.grid(color='white')
ax.xaxis.grid(color='white')

for ax, color in zip([ax], ['white', 'white', 'white', 'white']):
    plt.setp(ax.spines.values(), color=color)
    plt.setp([ax.get_xticklines(), ax.get_yticklines()], color='black')

width = 0.4

ind = np.arange(len(myx))

b1 = plt.bar(ind, nm, width, label='Nuoret miehet', color='royalblue')
b2 = plt.bar(ind, km, width, label='Työikäiset miehet', bottom=nm, color='cornflowerblue'
b3 = plt.bar(ind, vm, width, label='Eläkeikäiset miehet', bottom=nkm, color='lightsteelbl



b4 = plt.bar(ind + width, nn, width, label='Nuoret naiset', color='indianred')
b5 = plt.bar(ind + width, kn, width, label='Työikäiset naiset', bottom=nn, color='lightco
b6 = plt.bar(ind + width, vn, width, label='Eläkeikäiset naiset', bottom=nkn, color='mist


plt.xticks(ind + width / 2, labels, size=25, rotation=45)
plt.yticks(size=25)
plt.title('\n Vantaan kaupungin väestöennuste \n', fontsize=45)
plt.ylabel('\n Ihmisten määrä \n', fontsize=35)
plt.xlabel('\n Vuosi \n', fontsize=35)
plt.legend(fontsize=25)
plt.show()
```

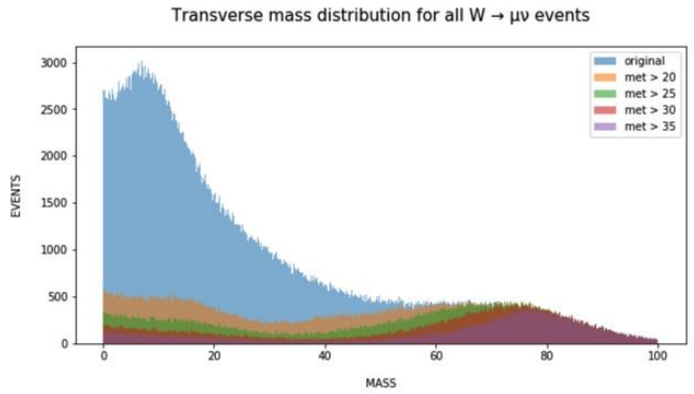
```
#set the size for the graph -- really important
fig = plt.figure(figsize=(20, 10))
ax = fig.add_subplot(1,1,1)
ax.set_axisbelow(True)
ax.yaxis.grid(color='gray', linestyle='dashed')
ax.xaxis.grid(color='gray', linestyle='dashed')

#This is not needed
#for i in ages:
#    ind = a.index[a.Age == i][0]
#    entries = list(a.iloc[ind])[2:]
#    plt.plot(myx, entries, label=i+1)


#bring created lists to the graph
p1 = plt.bar(myx, nuoret, label='Nuoret 1-15v')
p2 = plt.bar(myx, keski, label='Työikäiset 16-64v', bottom=nuoret)
p3 = plt.bar(myx, vanhat, label='Eläkeikäiset 65-99v', bottom=nk)
#plt.bar(myx, kaikki, label='Kaikki') #not needed in bar graphs


#visual adjustments -> improved visuality = easier to understand

plt.xticks(myx, labels, size = 15, rotation=45)
plt.yticks(size = 15)
plt.title('\n Vantaan koko kaupungin väestöennuste \n', fontsize= 23)
plt.xlabel('\n Vuosi \n', fontsize= 15)
plt.ylabel('\n Ihmisten määrä \n', fontsize= 15)
plt.legend(fontsize = 15)
plt.show()
```

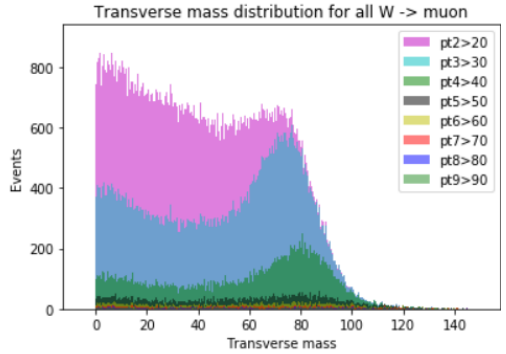


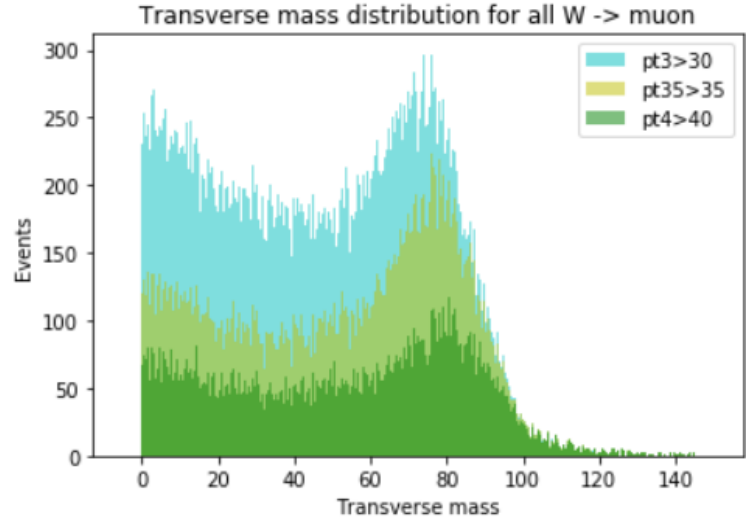

Vantaan koko kaupungin väestöennuste
Nuoret 1-15v
Työikäiset 16-64v
Eläkeikäiset 65-99v
Ihmisten määrä
0
50000
100000
150000
200000
250000
300000
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
Vuosi

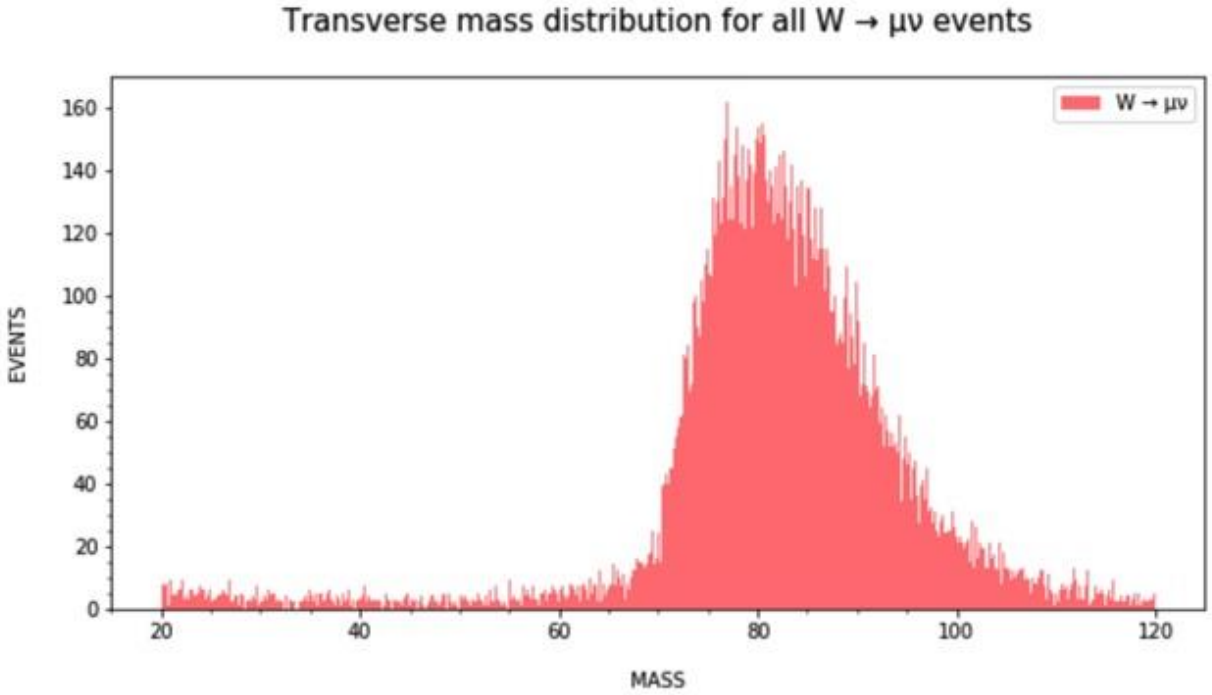
Vantaan kaupungin väestöennuste

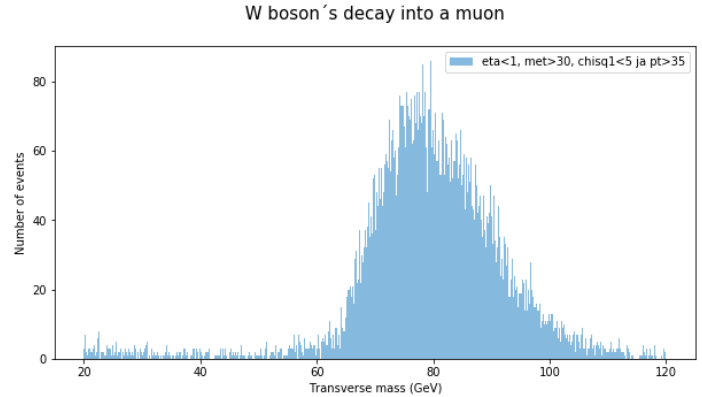


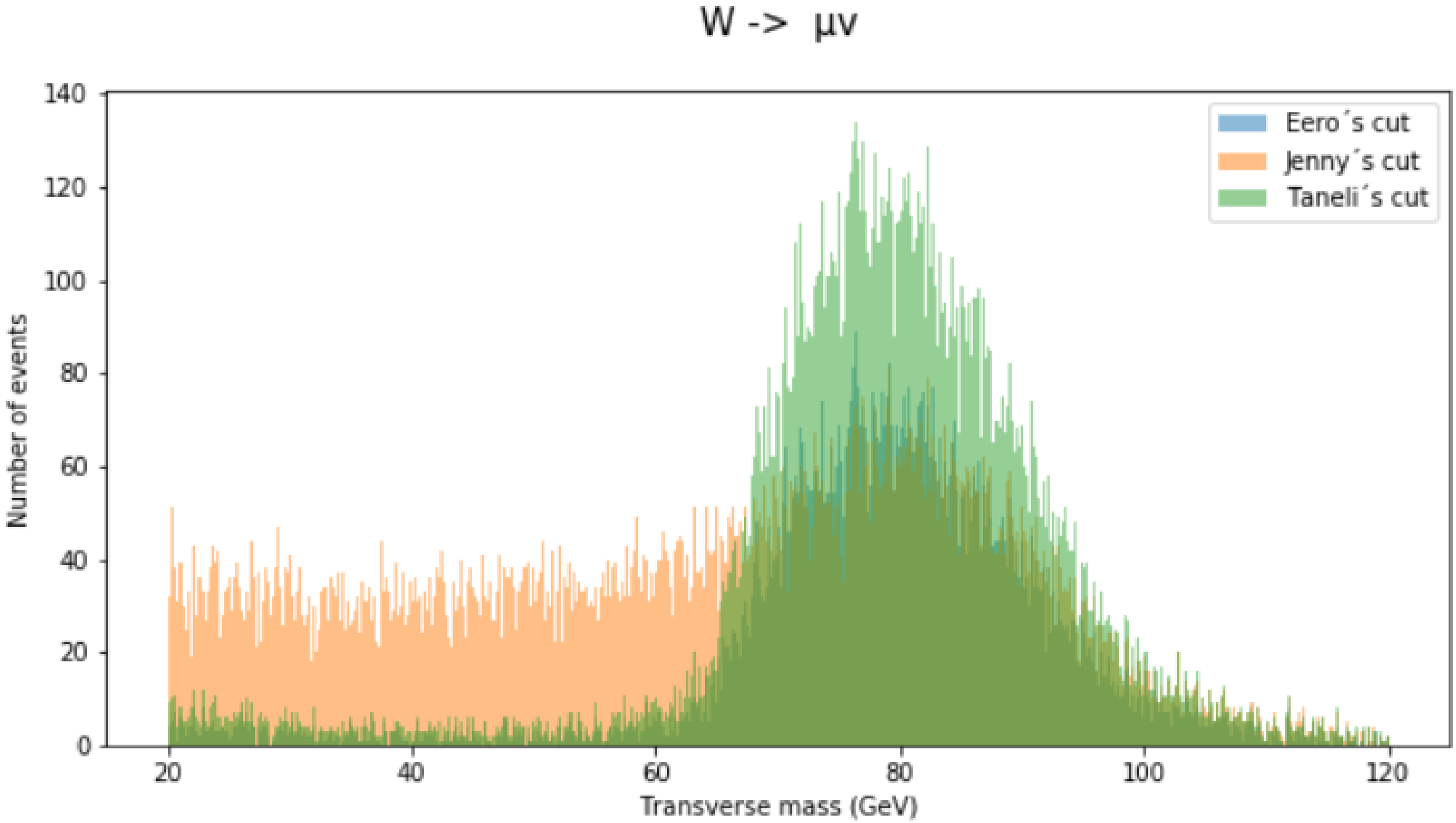

Transverse mass distribution for all W → μν events
original
met > 20
met > 25
met > 30
met > 35
EVENTS
MASS


Transverse mass distribution for all W -> muon
pt2>20
pt3>30
pt4>40
pt5>50
pt6>60
pt7>70
pt8>80
pt9>90
Events
Transverse mass


Transverse mass distribution for all W -> muon
pt3>30
pt35>35
pt4>40
Events
Transverse mass


Transverse mass distribution for all W → μν events
W → μν
EVENTS
MASS


W boson´s decay into a muon
eta<1, met>30, chisq1<5 ja pt>35
Number of events
Transverse mass (GeV)


W -> μν
Eero´s cut
Jenny´s cut
Taneli´s cut
Number of events
Transverse mass (GeV)
0
20
40
60
80
100
120
140
20
40
60
80
100
120

Thank you for your attention