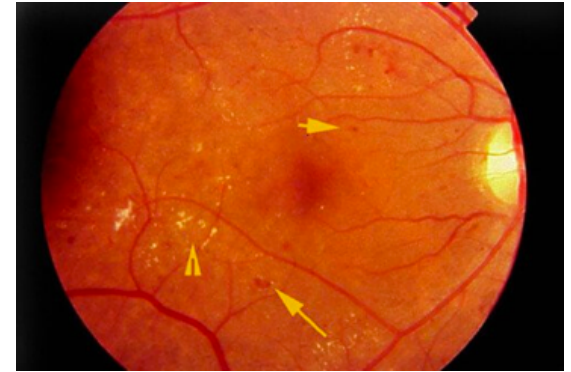# Announcements

- Readings (Week 7 due Sunday 2/23)
  - Ethics!

- **Assignment** 02
  - Due: Friday 2/21/2020, 11:59pm
  - Email me: 1 iPython notebook on Text **or/and** Audio (the whole notebook!)
  - Email me: any data (<5MB) -> otherwise put it in Google Drive and send me a link
  - Tell a story: describe what you are doing (within the notebook)
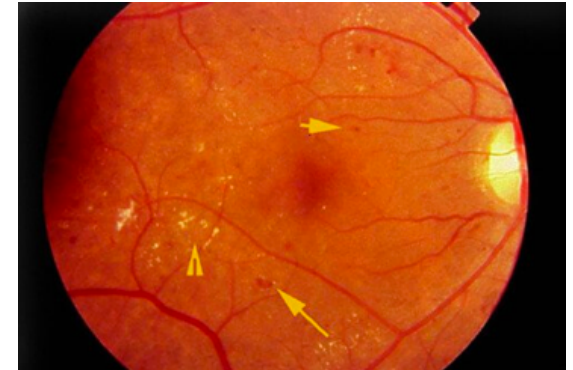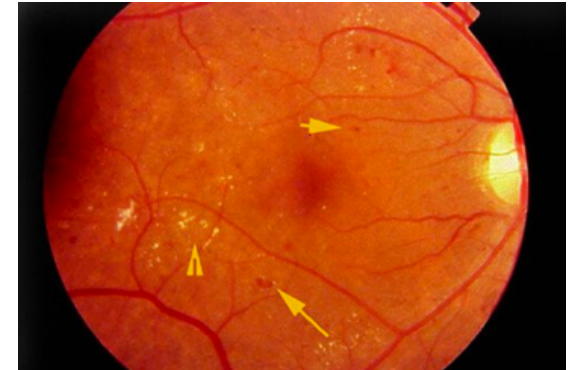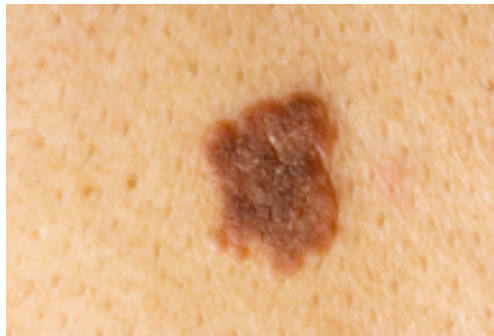
- **Today** Health Data
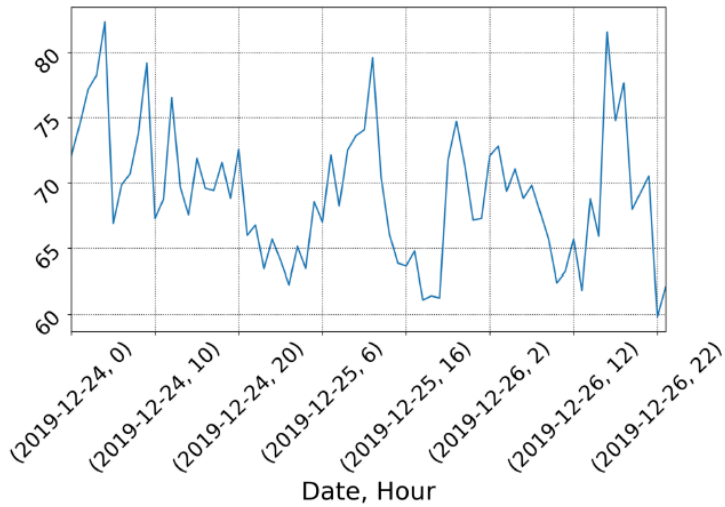 - How to?

What is Health Data?

What is Health Data?

What is Health Data?

What is Health Data?

What is Health Data?
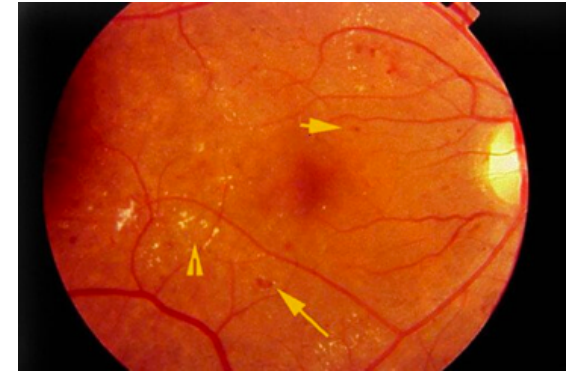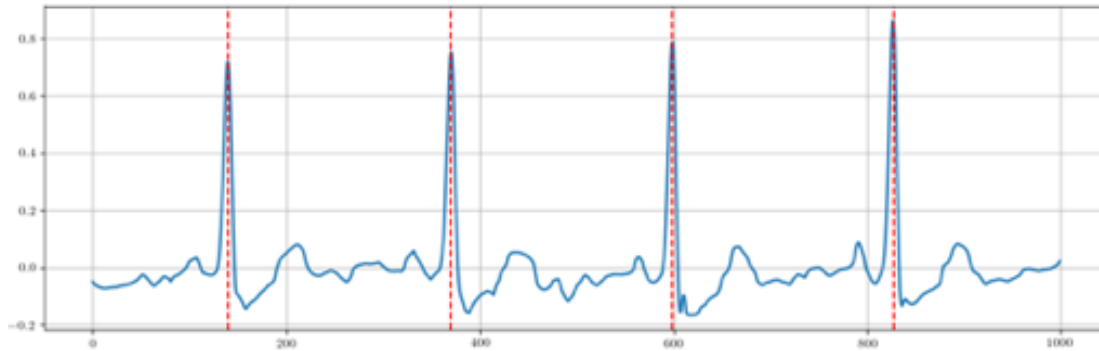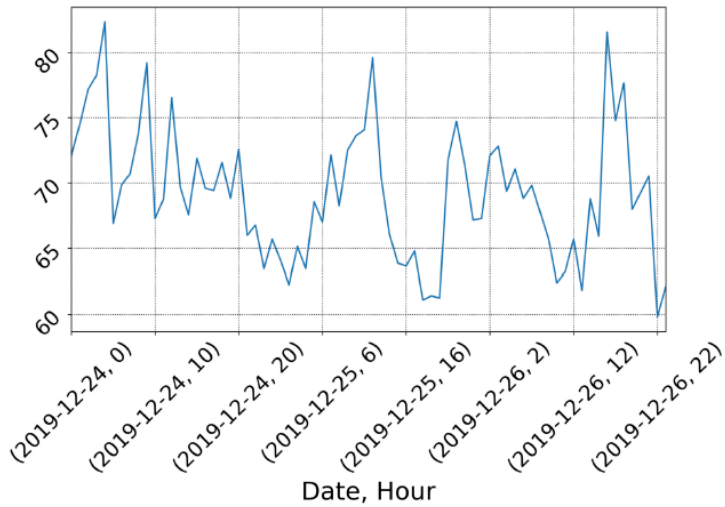
Augmenting diagnostic vision with AI

- ○ Radiologist -> 10M images
- ○ Dermatologist -> 200k skin lesions
- ○ Pathologist -> 100k specimens
- Deep Learning
  - ○ Diabetic retinopathy and macular edema from retinal fundus images = ophthalmologist
  - ○ Skin cancer = dermatologist
- Issues
  - ○ Labelled data
  - ○ Test with clinical trials
- Diagnosis?
  - ○ No explanatory power
  - ○ Support clinicians
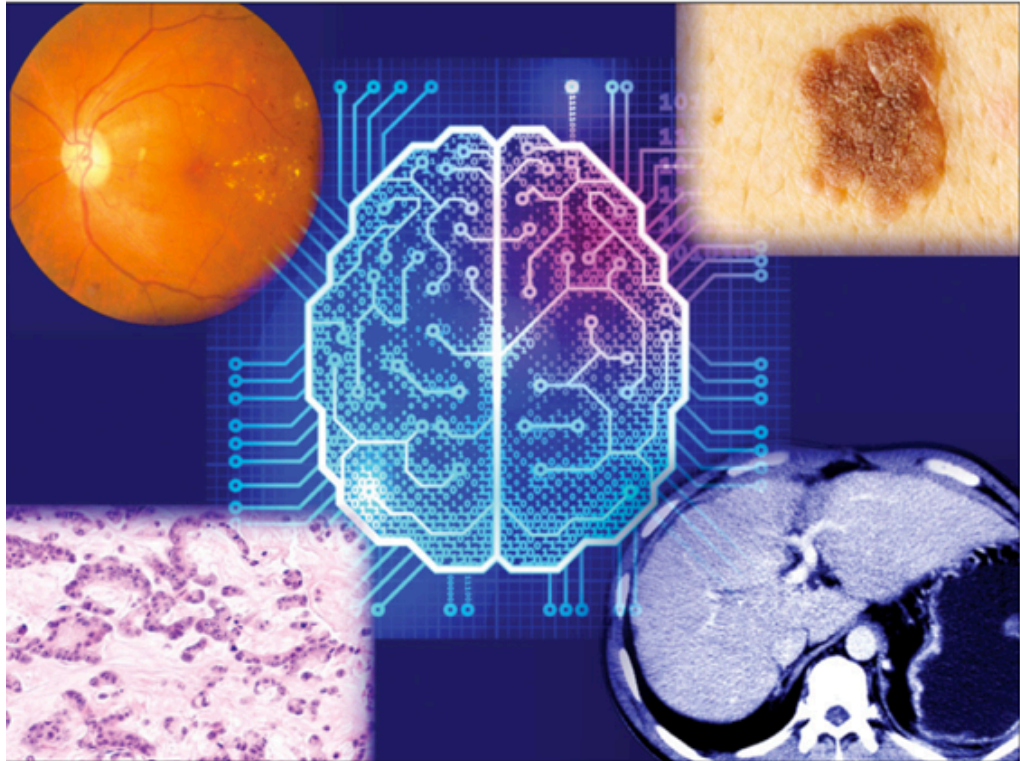
Wrist-worn: resting heart rate and sleep
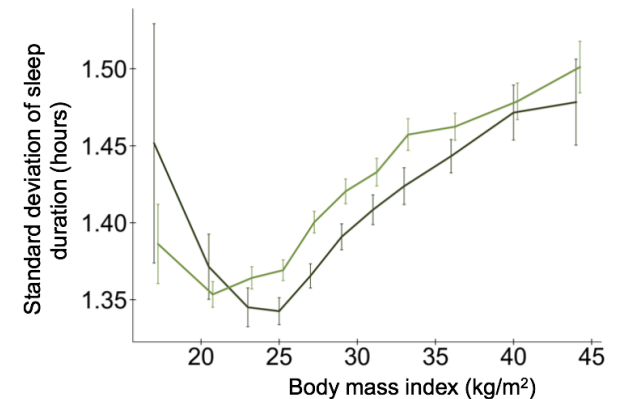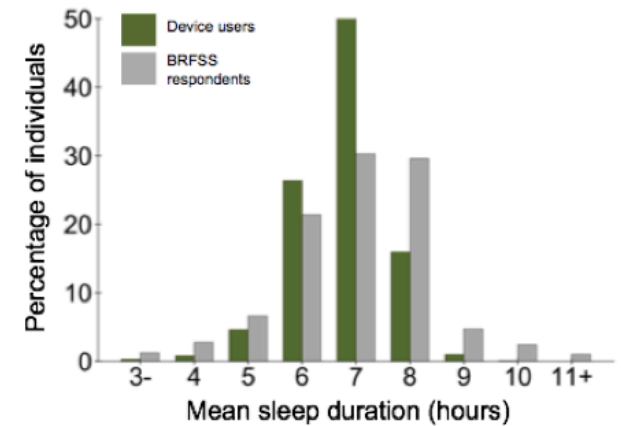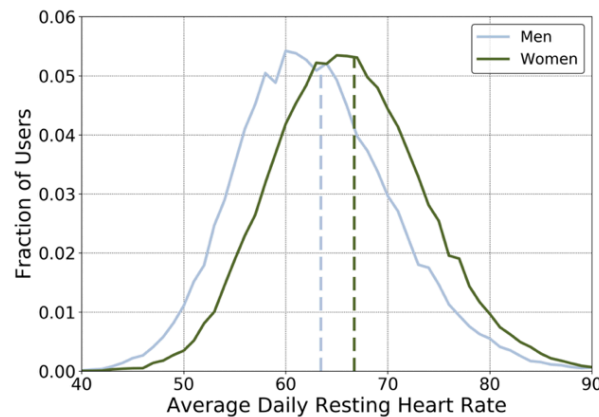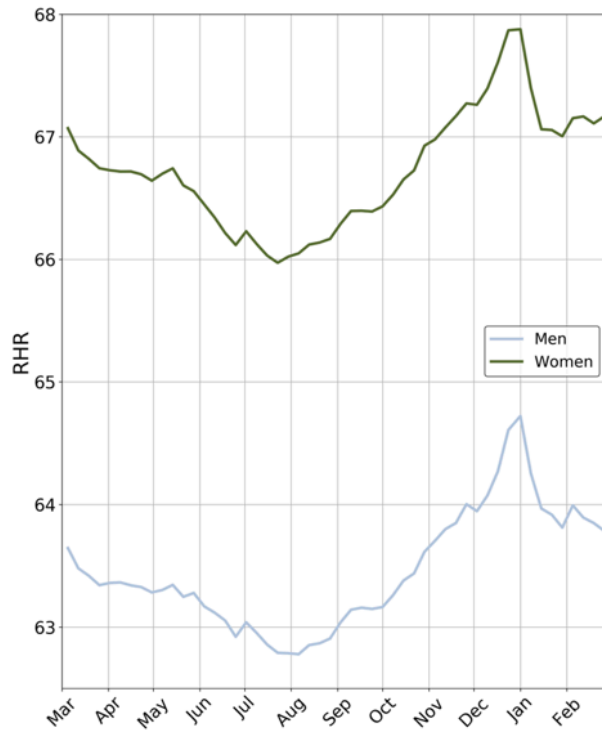
- 200,000 individuals over 2 years

- Normal daily RHR can vary by as much as 70 bpm between individuals
- Small but significant seasonal trends

- Over 30% of adults do not meet sleep guidelines (passive monitoring)
- Both sleep duration and variability contribute to overall health

Sleep data

- Total amount of sleep
- What about the different phases of sleep?
  - Hypnogram
    - Awake
    - Light
    - REM
    - Deep
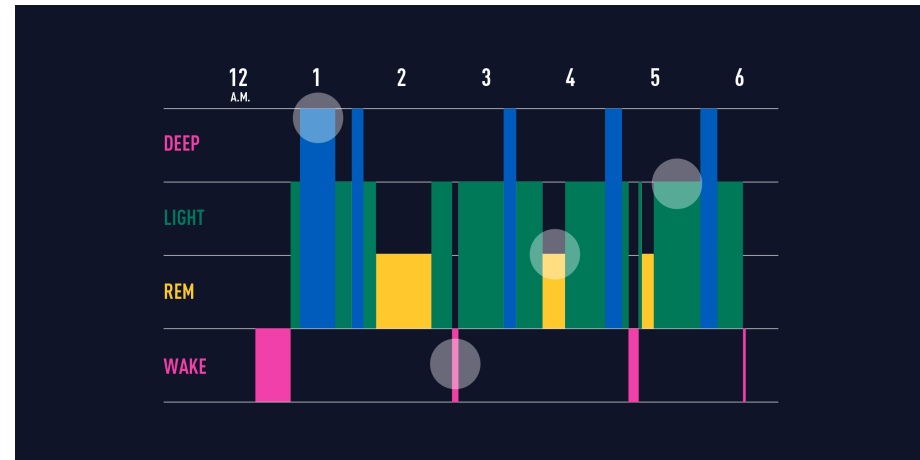
Sleep data

- Total amount of sleep
- What about the different phases of sleep?
  - Hypnogram
    - Awake
    - Light
    - REM
    - Deep

How does a Sleep file look like?

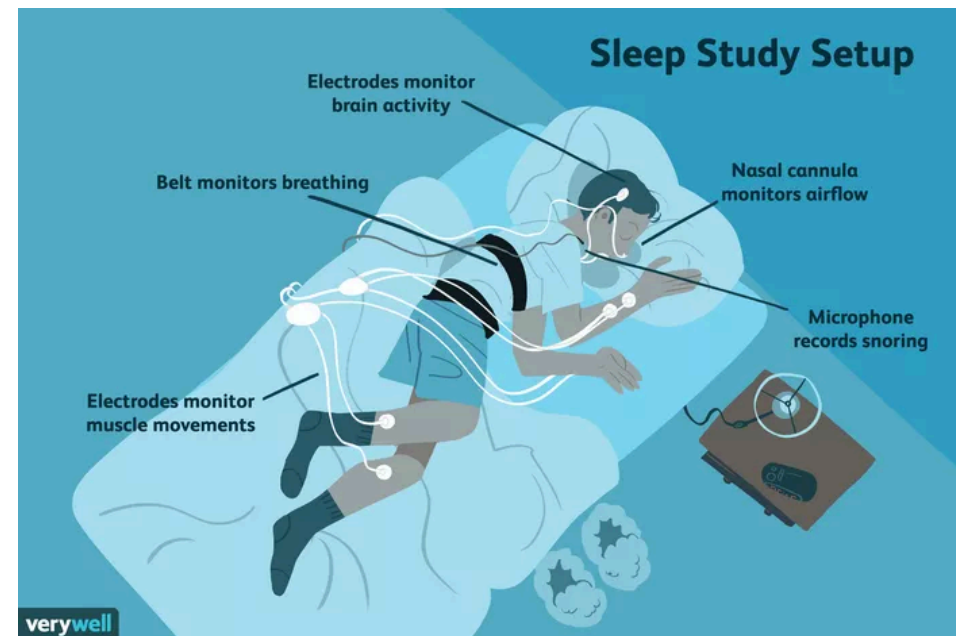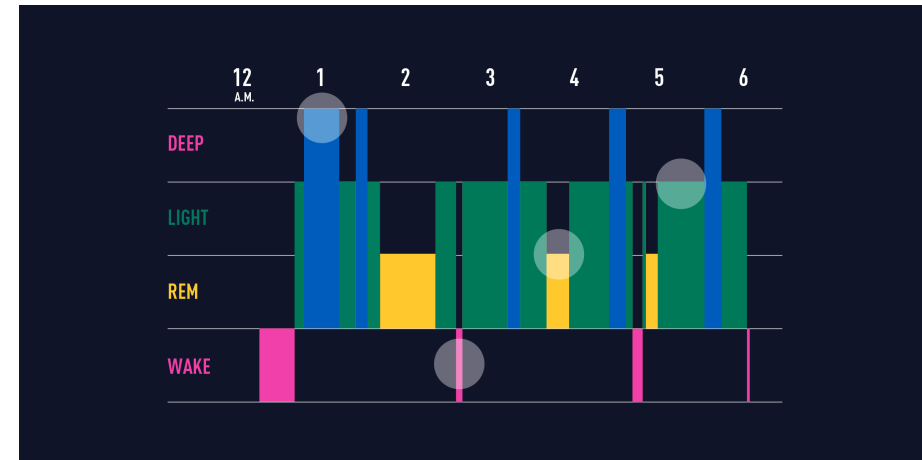| # Epoch Number	Stg	MicAr	{#}Any Apnea	{#}Apnea Obstructive	{#} Apnea Mixed	{#}Apnea Central	{#}Hypopnea Total	SUM(#DESAT_100_REM, {#}) Desaturation	{#}Leg Movement
79	W	0	0	0	0	0	0	0	0
80	W	0	0	0	0	0	0	0	0
81	W	0	0	0	0	0	0	0	0
82	W	0	0	0	0	0	0	0	0
83	W	0	0	0	0	0	0	0	0
84	W	0	0	0	0	0	0	0	0
85	N1	0	0	0	0	0	0	0	0
86	W	1	0	0	0	0	0	0	0
87	W	0	0	0	0	0	0	0	0
88	W	0	0	0	0	0	0	0	0
89	N1	1	0	0	0	0	0	0	2
90	W	1	0	0	0	0	0	0	1
91	N1	1	0	0	0	0	0	0	0
92	N1	1	0	0	0	0	0	0	0
93	N1	1	0	0	0	0	0	0	2
94	W	0	0	0	0	0	0	0	0
95	W	0	0	0	0	0	0	0	1
96	W	0	0	0	0	0	0	0	3
97	N1	1	0	0	0	0	0	1	0
98	N2	0	0	0	0	0	0	0	0
99	W	1	0	0	0	0	0	0	0
100	N1	0	0	0	0	0	0	0	0
101	N2	1	0	0	0	0	0	0	0
102	N2	1	0	0	0	0	0	0	0
103	N2	1	0	0	0	0	0	0	0
104	N2	1	0	0	0	0	0	0	0
105	N1	1	0	0	0	0	0	0	0
106	N2	1	0	0	0	0	0	0	0
107	N2	0	0	0	0	0	0	0	0
108	N2	0	0	0	0	0	0	0	0
109	N2	0	0	0	0	0	0	0	0
110	N2	0	0	0	0	0	0	0	0
111	N2	0	0	0	0	0	0	0	0
112	N2	0	0	0	0	0	0	0	0
113	N2	0	0	0	0	0	0	0	0
114	N3	0	0	0	0	0	0	0	0
115	N3	0	0	0	0	0	0	0	0
116	N3	0	0	0	0	0	0	0	0
117	N3	0	0	0	0	0	0	0	0
118	N3	0	0	0	0	0	0	0	0

What about a sensor data?

- JSON
  - JavaScript Object Notation
  - Object:
    - Unorder set of name/value pairs
  - Array
    - Ordered collection of values
  - Value
    - String
    - Number
    - Boolean
    - Object
    - Array

{'sleep': [{'dateOfSleep': '2020-01-17',
  'duration': 29460000,
  'efficiency': 96,
  'endTime': '2020-01-17T07:20:30.000',
  'infoCode': 0,
  'isMainSleep': True,
  'levels': {'data': [{'dateTime': '2020-01-16T23:09:30.000',
    'level': 'wake',
    'seconds': 570},
   {'dateTime': '2020-01-16T23:19:00.000',
    'level': 'light',
    'seconds': 1200},
   {'dateTime': '2020-01-16T23:39:00.000', 'level': 'deep', 'seconds': 1650},
   {'dateTime': '2020-01-17T00:06:30.000', 'level': 'light', 'seconds': 360},
   {'dateTime': '2020-01-17T00:12:30.000', 'level': 'rem', 'seconds': 300},
   {'dateTime': '2020-01-17T00:17:30.000',
    'level': 'light',
    'seconds': 1320},
   {'dateTime': '2020-01-17T00:39:30.000', 'level': 'deep', 'seconds': 1770},
   {'dateTime': '2020-01-17T01:09:00.000',
    'level': 'light',
    'seconds': 1920},
   {'dateTime': '2020-01-17T01:41:00.000', 'level': 'rem', 'seconds': 930},
   {'dateTime': '2020-01-17T01:56:30.000',
    'level': 'light',
    'seconds': 1440},
   {'dateTime': '2020-01-17T02:20:30.000', 'level': 'deep', 'seconds': 2190},
   {'dateTime': '2020-01-17T02:57:00.000', 'level': 'light', 'seconds': 810},
   {'dateTime': '2020-01-17T03:10:30.000', 'level': 'rem', 'seconds': 1320},
   {'dateTime': '2020-01-17T03:32:30.000',
    'level': 'light',
    'seconds': 1110},
   {'dateTime': '2020-01-17T03:51:00.000', 'level': 'deep', 'seconds': 420},
   {'dateTime': '2020-01-17T03:58:00.000',
    'level': 'light',
    'seconds': 1740},
   {'dateTime': '2020-01-17T04:27:00.000', 'level': 'rem', 'seconds': 3270},

Our dataset

- The participant spent the night from 15 to 16 of Jan in the sleep lab
- We have data from
  - Sensor 1
  - Sensor 2
  - Gold standard polysomnography (PSG)
- Goal
  - Compare the two sensors to the PSG
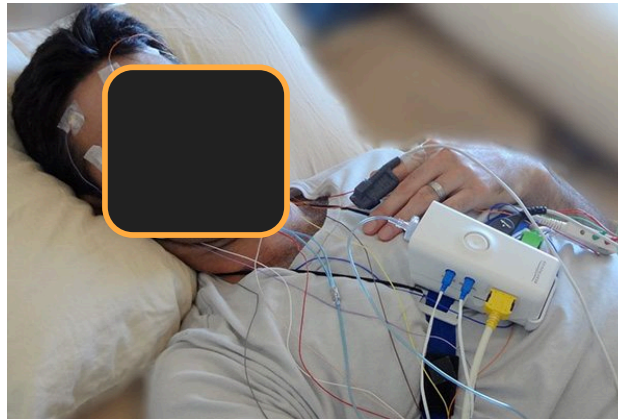
Today

- Sensor 1
    - JSON
    - 1 value of the hypnogram EVERY 5 minutes
    - The hypnogram is a string of characters

- Sensor 2
    - JSON
    - Precision of 30 seconds
    - One object with initial time and duration of each phase of sleep

- Gold standard
    - A table in html format
    - Precision of 30 seconds

- **Which sensor is… more accurate?**
    - Create an output with phases of sleep (from 11pm to 6am, 1 value every 30 sec)
    - Plot it
    - Compare it to the gold standard